Tracking referents

Taylor, Ryan Christopher

RYAN C. TAYLOR

TRACKING REFERENTS

Markedness, World Knowledge and Pronoun Resolution


university of groningen

CLCG
Center for Language and Cognition Groningen

bcn

The work presented here was carried out under the auspices of the Center for Language and Cognition Groningen (CLCG) of the Faculty of Arts of the University of Groningen and the School for Behavioural and Cognitive Neuroscience (BCN) of the University Medical Center Groningen.

Grodil Dissertations in Linguistics 116
ISSN: 0928-0030
ISBN: 978-90-367-6187-1 (Print edition)
ISBN: 978-90-367-6329-5 (Ebook)

Printed by: Wöhrmann print services, Zutphen, The Netherlands.
Cover design: Wilber Oswaldo Nucamendi Madrigal.
wilberoswaldo.mx@gmail.com – www.behance.net/wilberoswaldo

RIJKSUNIVERSITEIT GRONINGEN

TRACKING REFERENTS

MARKEDNESS, WORLD KNOWLEDGE AND PRONOUN RESOLUTION

Proefschrift

ter verkrijging van het doctoraat in de
Letteren
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
donderdag 4 juli 2013
om 11.00 uur

door

RYAN CHRISTOPHER TAYLOR

geboren op 7 juli 1983
te Edmonton, Canada

Promotor:	Prof. dr. G. Redeker
Copromotores:	Dr. J. C. J. Hoeks Dr. L. A. Stowe
Beoordelingscommissie:	Prof. dr. S. Baum Prof. dr. P. Hendriks Prof. dr. S. A. Kotz

The great tragedy of Science — the slaying of a beautiful hypothesis by an ugly fact.

— Thomas H. Huxley

There is a crack in everything, that's how the light gets in.

— Leonard Cohen

Dedicated to Velia del Pilar Altamira de Taylor,
for turning limes into, well, anything you could want.

Lo que no tenemos lo encontramos en un amigo. Creo en este obsequio y lo cultivo desde la infancia. No soy en ello diferente a la mayor parte de los seres humanos. La amistad es la gran liga inicial entre el hogar y el mundo. El hogar, feliz o infeliz, es el aula de nuestra sabiduría original pero la amistad es su prueba. —Carlos Fuentes

---

ACKNOWLEDGMENTS

I feel that I've been particularly lucky in the people I've met on the long road to getting the book that you're holding published.

John, Laurie, and Gisela, you've been unrelentingly supportive. I can't help but feel lucky to have had such an excellent group of supervisors. John, it's been a real pleasure to work with someone so serious about getting the stimuli and the methods right. They can even be a strange sort of poetry, or at least that's what I told myself when I was re-jigging a third draft after a second pretest on a Saturday afternoon. Gisela, I hope that I can someday reach your level of terminological precision in my writing. Our discussions have made me set the bar a lot higher. It's been a pleasure working with all of you and I hope there are more collaborations to come. Laurie, I'm not sure how you became a walking dictionary of psycholinguistic effects — not to mention fantasy literature, pre-1970's television, folk music, and cooking. Not many supervisors will zip-line at their student's wedding in the rain forest or let their students drop in to use the washing machine on a Sunday afternoon. The areas where you are one of a kind make for a long list. This project wouldn't have turned out as well as it has without the three of you.

To my reading committee, Shari Baum, Petra Hendriks and Sonja Kotz, I would like to express deep gratitude. Your criticisms have made this thesis a stronger piece of writing. It is an honour to have three scientists whose work I admire so much as members of the committee.

It's all too seldom that administrative personnel and other support staff are thanked in dissertations. When the wheels

of administration are whirring efficiently, the people behind it are perhaps easily overlooked. John Nerbonne was chair of the CLCG throughout most of my time in Groningen, and made many comments that were helpful from both a scholarly and project management perspective. Wyke van der Meer was always supportive of my project and I'm very thankful for her help. I would also like to extend thanks to Diana Koopmans and Janine Wieringa of BCN. The *audiovisuele dienst* of the university was able to supply me with whatever recording equipment I needed, working perfectly, at the drop of a hat. Peter Albronda was very helpful in training both me and those helping me in the set-up of ERP, and again, everything always ran smoothly.

Thank you to Vânia de Aguiar, Agnes Engbersen, Sandra Hanekamp, Jakolien den Hollander, Trudy Krajenbrink, Neri Kamphuis, Kimon Lèfas, Hanneke Loerts, Fernando Nieto Morales, Rocío Pineda Duran and Adriá Rofes for help in developing stimuli and testing participants. I would also like to thank Laura Kertz for providing her materials. I would like to thank Universidad Iberoamericana, especially Alejandra Santoyo, for help in recruiting participants.

Thank you to Susanne Borgwaldt, Gregg Moor and Norma Moor for their help with editing and proof-reading over the years.

I've have benefitted greatly from the collegiality in the department, from participation in discussion groups to hallway chats, I'd like to thank: Veerle Baaijen, Karin Beijering, Ildikó Berzlánovich, Laura Bos, Harm Brouwer, Susanne Brouwer, Gisi Canizzaro,Tim van de Cruys, Diana Dimitrova, Christina Englert, Nanna Haug Hilton, Jack Hoeksema, Bart Hollebrandse, Peter Hendriks, Charlotte Koster, Gideon Kotze, Dörte de Kok, Harrie Mazeland, Ben Maassen, Ruggero Montalto, Jacolien van Rij, Pavel Rudnev, Margreet van Koert, Nynke van der Vliet, Martijn Wieling. In particular I'd like to thank Trevor Benjamin and Kashmiri Stec for being my paranymphs, for being such great friends, and for helping to take some of the weight of the defence off my shoulders.

In Groningen, I'm also leaving behind a lot of friends who weren't at the university with me. Lisa, Viv, Geoff, Dave, Anne, Auke, Gunn, Lammertjan, Jaimy, Dennis, Lara, Elena, Thomas, Belen, Sylvain, Bruno, Paul, Fernando – let's see each other

again soon and often. Romke Dam, Lorina Gjonaj and Elena Melillo have at various times let me stay at their houses for a prolonged period, for which I am particularly grateful.

Inevitably there are people I have neglected to mention. I am sorry. In the years it has taken to write this I have had support from many people in many ways, and I am grateful for it all. Would a beer/coffee make up for it?

Doña Pilar y Don Armando, siempre me han hecho sentir mucho más que bienvenido y ahora puedo decir que realmente comprendo que significa familia Mexicana. Muchas gracias también a Daniel y a Santiago por hacer mi estancia en México tan agradable.

Mom and Dad, Medical doctors and education researchers both give the same advice: Choose your parents carefully. Thank you for always encouraging me to study hard. I sometimes wonder whether I (being true to form) took your advice further than you intended. Nevertheless, it's largely your fault I'm here. I am also grateful to my Grandparents for the many hours of good conversation and advice over the years.

Mark, It's amazing how much a brother can grow on you. I used to often wish that you would take a hike. Now I often wish that we could take a hike.

And Velia, you have given me constant, unwavering support. They say that behind every great man stands a great woman. It turns out that great women are behind men with an oddly persistent interest in Language and the Human mind too.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

BRIEF–A Behavioural Rating Inventory of Executive Function–Adult self-report

ERP Event-Related Potential

EEG	ElectroEncephaloGram
fMRI	functional Magnetic Resonance Imaging
PFS	Parallel Function Strategy
ROI	Region of interest
SAS	Subject Assignment Strategy
SPR	Sefl-Paced Reading
SOV	Subject Object Verb
SVO	Subject Verb Object

1

INTRODUCTION

A central feature of language is that it allows us to convey information about things, what happens to them, and the relations between them (Jakobson, 1960). This function of language is facilitated by the use of anaphoric expressions, which allow speakers to avoid repeating long and monotonous expressions like *the little red-headed girl from the peanuts comic strips*. In English or Dutch, entities or characters that have already been mentioned (as in 1) or are easily understood from context (as in 2) can be referred to with pronouns,

(1) *I talked to Terry. He says hello.*

(2) *I went to the pharmacy. He filled my prescription very quickly.*

In (1), Terry was just mentioned, so it's trivial to understand who the pronoun refers to. In (2), the context of a pharmacy makes a pharmacist salient enough that the referent for *he* can be inferred.

Although using short referring expressions like pronouns allows speakers and writers to avoid arduously repeating themselves, at the same time, use of these shorter referring expressions may lead to comprehension difficulties for readers and listeners, since these short referring expressions are necessarily more ambiguous. Whereas *the little red-headed girl from the peanuts comic strips* refers to only one individual, pronouns can be used to refer to different individuals and take their meaning from context. However, keeping track of who is being referred to from a comprehender's point of view is not an easy task. In fact, the evidence suggests that comprehenders must often use multiple sources of information in order to understand who a pronoun refers to. It is possible that not all of these informa-

tion sources are used by the processor in all instances, and that the extent to which they are important for the resolution process may vary by language (Morrett and MacWhinney, 2013), but there is strong evidence suggesting that most of the time, multiple factors work in concert.

Past research (e.g. Morrett and MacWhinney, 2013; Järvikivi et al., 2005) into pronoun comprehension has focused to a large extent on structural factors such as grammatical role (of both referent and pronoun), order-of-mention (participants mentioned first in a sentence are more accessible than participants mentioned later) and clause recency (concepts mentioned in the most recent clause are more accessible than concepts mentioned in an earlier clause). Research has put less emphasis on the more conceptual aspects of pronoun comprehension. As a consequence, not much is known on how *conceptual information* is integrated in the comprehension of pronouns, especially where on-line processing is concerned. So research in that area is definitely needed. In addition, within psycholinguistics there has been burgeoning interest in *information structure*, which describes the principled ways in which languages deal with the difference between hearer new and old information and discourse new and old information (e.g. Lambrecht, 1996; Prince, 1981; Ward et al., n.d.).

1.1 CONCEPTUAL INFORMATION

In some accounts of pronoun resolution, conceptual information, in particular *world knowledge* or *plausibility*, is the most important determinant. Under these accounts, language users will always eventually choose the most plausible antecedent, regardless of other factors (e.g. Huang, 2000). On the other hand, it seems that language must also offer the opportunity to describe an event in the real world that is not plausible at all (e.g. *John witnessed Bill rob a bank, but then* JOHN *was arrested by mistake*). Under that view (c.f. Kehler, 2002), less plausible antecedents can become preferred, but their unlikeliness must be indicated by some linguistic means, such as accentuation (or other means of adding emphasis).

Although most researchers have assumed that plausibility will have some effect, they have paid relatively little attention to it as a contributing factor. Moreover, some of the earlier studies on pronoun resolution do not appear to have adequately controlled for the effects of world knowledge. For example, in testing for plausibility, Crawley et al. (1990) asked judges whether both names were plausible antecedents. If the question was indeed so posed, it seems likely to have led the judges to say yes, unless the plausibility differences were very large. Chambers and Smyth (1998) seem not to have controlled for plausibility with any measures beyond the authors' intuition either. This suggests that the differences in outcome may be due to the plausibility of different readings across experiments, rather than the factors that were being investigated. It thus seems to be necessary to examine this issue more closely. I will do this in a set of self-paced reading experiments described in Chapter 2, in a series of off-line antecedent selection studies in Chapter 3, and in an ERP study presented in Chapter 4.

1.2 INFORMATION STRUCTURE

A factor that is presumably as important as conceptual information is information structure. Information structure indicates what is given information and what is new information in a specific utterance and is reflected in the actual utterance in terms of word order, syntactic structure and/or prosodic realization of the utterance. Languages have been shown to differ in terms of the options they have at their disposal (Lambrecht, 1996). Hearers can make use of these information structural cues to connect the incoming information with the knowledge they have about the current discourse and the world in general. Several theorists have suggested that a speaker's choice of syntax or intonation results from what the *hearer* could be assumed to know (Pierrehumbert and Hirschberg, 1990; Steedman, 2007; Ladd, 2008; Ward et al., n.d.). Syntax and prosody are thought to be chosen in such a way that new information is marked, and old information unmarked. In general, given information is positioned first in an utterance, and new information last (Ward et al., n.d.; among many others). Order-of-mention (which I control for in

Chapter 2, and examine in Chapter 3) has been found to be extremely important in pronoun comprehension (Smyth, 1994; Kehler et al., 2008, inter alia).

Prosody, and perhaps most importantly *accent*, can also be co-opted for marking information, in languages like English and Dutch. There is still argument about the exact form a theory of accent should take, in particular because speakers are inconsistent in how they produce accent (Breen et al., 2010). Nevertheless, there is growing consensus that speakers will accent new information, as in example 3 (taken from Breen et al., 2010).[1]

(3) a. *Who fried an omelet this morning?*
b. DAMON *fried an omelet this morning.*

Moreover, Breen et al. present evidence that some of the inconsistencies are due to speakers not being aware of ambiguities, since speakers produced disambiguating accent when it was pointed out that an ambiguity existed.

Lending credence to this analysis, Dimitrova et al. (2009) found that speakers will accent information that is surprising to them—even if the information is contextually given (e.g. the colour *blue* in a context with all blue objects), and it is only the combination that is surprising (e.g. *blue banana*). In fact, speakers with better ability to direct processing resources (executive function) were the most likely to accent unexpected information. Thus the inconsistent empirical results may be the result of varying speaker awareness rather than a problem with the linguistic analysis. Although the degree to which discourse new/given, hearer new/-given and contrast are represented in intonation is uncertain, there is agreement that some sort of distinction between given and new exists (Breen et al., 2010). The generalization that new information is accented is captured by Huang's application of neo-Gricean theory to anaphora: A marked (rarer) referring expression tends to refer to something more difficult to retrieve (Huang, 1991).

1 Throughout the thesis accented words are indicated with SMALL CAPS.

1.3 INVESTIGATING MODELS OF PRONOUN RESOLUTION

It is still unclear what psychological processes the hearer uses to decipher the reference, or what model provides the best description of pronoun processing. In order to understand the contribution of information structure and world knowledge in pronoun resolution, I have looked at how world knowledge affects pronoun comprehension in on-line self-paced reading (Chapter 2), I investigated how information structure and world knowledge lead to the choice of one referent over another using an antecedent selection task (Chapter 3), and I measured ERPs to investigate how information structure signals and world knowledge are integrated on-line (Chapter 4).

One of the most important ways of examining comprehension is to look at comprehension difficulty when resolving the pronoun does not go as planned. As yet there is no agreement on how to model the processes underlying comprehension difficulty. In this thesis I compare instances of the two major types of model: a leading serial model to a leading parallel model, using self-paced reading (Chapter 2). The most salient difference between these models is that parallel models assume that all information is used immediately, whereas serial models assume that some (often structural) information is used before other (often non-structural) information.

The particular serial parser that I examine, the Unrestricted Race Model (e.g. Van Gompel et al., 2000), assumes that ambiguity leads to *less* difficulty when it is first encountered, but more difficulty at a later point if the utterance is disambiguated. The parallel parser I examine, the Constraint-Based Model (e.g. McRae et al., 1998), predicts that ambiguity will often lead to difficulty.

I compare these models using self-paced reading, the workhorse of psycholinguistics. It involves a participant reading a sentence by pushing a button to advance from one word to the next on a screen, and timing the speed of reaction. Reactions at critical points are then compared, in order to distinguish between theoretical predictions. Results of a study by Swets et al. (2008) suggests that participants only process the stimuli deeply enough to complete the task, as described by the Good-Enough Processing Hypothesis (e.g. Ferreira and Patson, 2007). If this

is true, the finding that the results across studies have been disparate may be attributable to the task the participants performed while reading. Thus, I included task as a condition to see whether difference of task indeed leads to differences in results.

Modelling who a pronoun refers to is separate from which model best describes difficulty in comprehending a pronoun. However, like in comprehension difficulty, there is some evidence that world knowledge and markedness are combined in the choice of referent. Accordingly, most current models of referent choice either suggest that multiple sources of information are combined with a particular set of weightings in order to come up with an end result, or that it is the result of some ranking mechanism. There are two pertinent issues in the choice of referent. One is how on-line models can be applied to the question of choice. The second issue is parallel to how the models can be applied: do all sources of information contribute to referent choice, even if they contribute unequally? Or, does the highest ranked source of information block out all others? In terms of on-line models, the Unrestricted Race Model is not specific with regards to whether a ranking or weighting mechanism underlies choice of referent, and whether there is in fact interaction, or only one factor at a time is involved in the choice (for example, if world knowledge were to always over-rule markedness when there is a world-knowledge bias). The Constraint-Based Model, however, is quite specific about how choice of referent occurs. It suggests that it is effectively a process of weighting: Over time, sources of information are taken into account, and preference moves toward one or the other referent. When preference is sampled, it is effectively the weighting that is being sampled. Other approaches (e.g. Hoeks and Hendriks, 2011) make use of ranked constraints, as opposed to weighted sources of information.

The chief goal of this thesis is to understand how information structure and world knowledge interact. Thus, an important question is whether the workings of the processor result in an all or nothing choice, or a graded choice of referent, like the Constraint-Based Model described in Spivey (2008) or the optimality-theory based model described in Hoeks and Hendriks (2011). An all or nothing model would suggest that some sources

of information (such as plausibility, markedness or syntax) would determine referent choice whenever there was a bias. Thus if plausibility were highest-ranked, and there were some plausibility bias, the model would always choose the referent that is most plausible, and no other. Other factors would make themselves apparent when both referents were equally plausible. A graded choice would mean that different information sources each contribute somewhat to the choice of referent.

Some preliminary research has looked at accented pronouns (markedness) or at world knowledge, but not both together. Work on accented pronouns has suggested that accent causes hearers to choose an initially less active referent (Venditti et al., 2001), already suggesting that the processor takes several sources of information into account. In order to fully understand how these sources of information interact, an experiment directly probing their interaction is necessary, such as the antecedent choice experiment reported in Chapter 3.

1.3.1 *Neurolinguistic Modelling of Comprehension*

In addition to examining how world knowledge affects the speed of comprehension and referent choice, there is a further matter of how comprehension of these sources of information affects the brain. One method that is highly effective at linking psycholinguistics to neurolinguistics is ERP (Event-Related brain Potentials). Like self-paced reading, ERP studies have the advantage of a large body of literature going back over several decades. ERP has been used to great effect in the study of accent (one means of making something marked), semantic anomaly (related to world knowledge) and pronoun ambiguity. Furthermore, ERP is particularly well suited to the study of *spoken* language comprehension, since, unlike self-paced reading, it requires no response on the part of the listener. Comprehension processes underlying normal, fluid speech presented auditorily are easily captured.

Registration of ElectroEncephaloGram (EEG) involves an electrode cap placed on the scalp of the participant. The electrodes on the scalp pick up changes in electrical potential caused by the activity of large assemblies of neuronal cells (Luck, 2005). In ERP

experiments, the EEG signal is averaged at critical points where the stimuli differ, in order to separate the signal from the noise (which is assumed to be averaged out). ERP registration has excellent time resolution, and is able to pick up even very fast changes in electrical potential in the cortex. Additionally, ERP recording gathers information about *where* on the scalp these changes are taking place. The mapping of the scalp topography of a signal to localization is problematic, however, although it does allow researchers to reason about whether an effect generated in one condition or experiment has the same topography as an effect in another condition/experiment.

In Chapter 4, I look at specific components with known properties to see the effect of markedness and world knowledge on the processing of a pronoun, following the philosophy of Luck (2005). Chapter 4 therefore targets theories of what a specific component is expected to do, as opposed to overall neurolinguistic models of what components are expected to do. Past experiments have examined pronouns, markedness and world knowledge, although none have examined all three. The most prominent ERP experiments on pronouns have investigated ambiguity. Studies have found that ambiguity creates a component referred to as an NRef, a sustained frontal negativity (Van Berkum et al., 2007). However, since the pronouns I investigate are not ambiguous, it is interesting to find out whether this research is applicable. ERP research on markedness (accent) suggests that accented words provoke an *accent positivity* relative to unaccented words (Dimitrova, 2012). This has been taken to mean that accent causes increased attention to the accented word, and perhaps intensified processing, although it is uncertain how this would function on a pronoun, especially for unexpected referents.

In addition to early positivities, the experiment allows us to examine whether expectedness and accent have any effect on the N400 (a negative deflection in the ERP signal that is maximal around 400 ms after onset of the evoking stimulus), and the P600 (a positive deflection in the ERP signal that is maximal around 600 ms after onset of the evoking stimulus). The amplitude of the N400 is sensitive to the semantic fit of a word with respect to the preceding context. The P600, on the other hand, has received various interpretations, ranging from syntactic re-

vision to pragmatic processing (see Brouwer et al., 2012, for an overview).

1.4 OVERVIEW OF THE THESIS

Each of the chapters of this thesis examines aspects of how specific types of information are combined in pronoun resolution. In Chapter 2, I compare the results of a self-paced reading study with the predictions of different models of language comprehension that describe how ambiguity is resolved under these conditions in comprehension. I examine a parallel processing model (Constraint-Based Model; McRae et al., 1998) and a serial processing model (Unrestricted Race Model; Van Gompel et al., 2000). These models provide worked-out mechanisms that allow a researcher to take a linguistic stimulus, and from that stimulus create predictions about processing speed for a given word or passage. I also examine the Good-Enough Processing Hypothesis (Ferreira and Patson, 2007).

In Chapter 3, I look at world knowledge and markedness in English and in Spanish, examining the importance of order-of-mention and grammatical roles along the way, by asking participants to identify the referent of the pronoun. In both English and Spanish I look at how the markedness of the form used interacts with structural biases on the one hand and world knowledge on the other in the final interpretation of a pronoun.

In Chapter 4, I expand on the on-line dynamics of information use in the comprehension of pronouns, looking at both markedness and world knowledge. Using ERP allows me to comment on when each type of information begins to affect processing to a very accurate time resolution, and even allows comparison of how the input affects specific components of ERP response. Following on past research that has found evidence that unexpected or difficult constituents are accented (Dimitrova et al., 2009), I investigated the effects of accentuation and unexpectedness on pronoun resolution. In one condition, sentences contained a pronoun with an unexpected referent, since they described a situation that went against world knowledge (e.g. *someone* loses *a tournament and is congratulated*), in the control condition, the referent was expected (e.g. *someone* wins *a tourna-*

ment and is congratulated). By using ERP registration, I looked into the processing of unexpectedness, and at the effect of accenting on that process.

I conclude by relating the results to each other in Chapter 5, discussing what restrictions our results prescribe for models of pronoun resolution and suggesting areas for future research.

2

ENCOUNTERING AN AMBIGUOUS PRONOUN

Evidence for Competition in Referential Processing

ABSTRACT

Research on pronoun resolution has not yet produced a comprehensive model of referential processing. In this chapter, we investigate whether two well-known models of syntactic ambiguity resolution, the parallel Constraint-Based Model and the serial Unrestricted Race Model, can be successfully applied to pronoun resolution. In addition to these two models, we will consider a processing hypothesis as well, the Good-Enough Processing Hypothesis, which assumes that processing is to some extent strategic, and may depend on task requirements (we term this a hypothesis rather than a model since it pertains to the *depth* of processing, but does not offer a description of processing *per se*). Our participants read stories containing ambiguous subject pronouns following three task instructions: answer a question to disambiguate the referent of a pronoun, give a plausibility rating concerning the story as a whole, or simply read the stories attentively without any subsequent questions. Our results show that in the Disambiguation Question task, ambiguous pronouns were read more slowly than unambiguous pronouns, as predicted by the Constraint-Based Model. In the Plausibility Rating task, we found a similar, but smaller ambiguity disadvantage. When participants did not have to answer questions, there were no reading time differences at any of the regions we investigated. These results suggest that referential ambiguities give rise to competition between alternatives, but lower engagement tasks resulted in weaker results or null results, which the Constraint-Based Model cannot account for without modification.

2.1 INTRODUCTION

The processing of ambiguity in language has proven to be very important, because it speaks to the hotly contested problem of when information becomes available for use by the language processor. Is all information—morphology, syntax, semantics, discourse structure, pragmatics—used in parallel? Or is some information (e.g. syntax, topic structure) available before other information or perhaps given more weight? By studying ambiguity resolution, it is possible to find out what types of information are more strongly weighted. For this reason, both referential ambiguities and syntactic ambiguities have seen a lot of attention. In the experiment reported here we examine how well two models of ambiguity resolution are able to predict reaction times during the comprehension of referential ambiguities.

Below we examine two of the most current models of ambiguity resolution, looking at a serial parser, which examines only one option at a time, and a parallel parser, which examines all options in parallel. Following on the rationale of Swets et al. (2008) we hypothesized that differences between these models can be explained by the fact that self-paced reading experiments described in the literature frequently differ strongly in task. Some of these studies found evidence in favour of a parallel parser, whereas other studies found evidence against, hence it is possible that differences in task underlay discrepancies in one study relative to another. Since testing these models against each other requires an initial ambiguity that is later resolved, we work through an example sentence from the experiment showing what types of biases comprehenders encounter on-line. We then show how the different models predict that comprehenders deal with these biases, and the effect they will have on processing.

2.1.1 *Ambiguity and Processing*

Pronouns are a hotbed of potential ambiguity, and pronominal ambiguities can cause significant problems in comprehension. Take for example (1):

(1) *Frank played cards with Steve for a huge sum.*
Steve lost to Frank.
He ...

When the comprehender reaches the pronoun, *Frank* and *Steve* are both good continuations. Many factors have been shown to be involved in resolving pronominal ambiguity. In this paper, we will not concern ourselves with these specific factors, but rather focus on finding out what kind of mechanism can underlie the use of relevant constraints in the resolution process. At present, there is no elaborated model of how constraints are handled in pronoun processing. As a point of departure, we take existing models of syntactic ambiguity resolution and see whether they can be fruitfully applied to pronoun resolution as well.

2.1.2 *Models of Syntactic Ambiguity Resolution*

2.1.2.1 *The Constraint-Based Model*

A very influential model of syntactic ambiguity resolution is the Constraint-Based Model, in which all different kinds of constraints interact in a parallel process through which the syntactic option that receives the most support from the constraints is chosen (Trueswell and Tanenhaus, 1994; MacDonald et al., 1992; MacDonald, 1994; MacDonald et al., 1994). A computational implementation of this type of Constraint-Based processing is the well-known Competition-Integration Model (McRae et al., 1998). In this model, different sources of information are represented as *input* nodes. In the domain of referential processing, for example, there could be an *animacy* input node, since *animacy* has been shown to be involved in referent choice. In addition, there could be nodes representing *order-of-mention*, *recency*, *world knowledge bias*, etc.

Model simulation proceeds in a number of steps. First, activation from the input nodes is fed into *representation* nodes that each represent the possible referential solutions. This information is then fed back to the input nodes, after which the cycle is repeated until the difference in activation between the representation nodes has become high enough.

If there is a strong preference for one of the alternatives, for instance for the first-mentioned noun, *Frank*, processing would proceed quickly. Close *equibiases*, however, where both *Frank* and *Steve* are equally likely as referents, will lead to longer processing times. For example, at *He* the processor is uncertain who the pronoun refers to, since there may be a first-mention bias toward *Frank*, but a recency bias toward *Steve*. If, for example, there were a strong bias toward *Frank*, processing would proceed quickly. However, when the biases toward both names are roughly equal, the activation nodes each have highly similar activation levels—leading to more cycles of competition and longer processing times.

If the characters were *Frank* and *Sarah* instead of *Frank* and *Steve*, as in (2), it would be easy to resolve the pronoun based on gender, so a minimal number of cycles of competition would be required, leading to minimal processing difficulty and hence relatively short processing times.

(2) *Frank played cards with Sarah for a huge sum.*
Sarah lost to Frank.
He . . .

If it turns out later in the sentence that the processor has made the wrong decision, a round of competition will again be necessary. Look, for example, at the continuation in (3):

(3) *Frank played cards with Steve for a huge sum.*
Steve lost to Frank.
He left with a big smile.

Since smile is biased to *Frank* (the winner), the model is able to resolve the pronoun. However, since some of the time the model will have resolved toward *Steve* at the cutoff point, it will take several cycles of competition to sort this out, leading to a relatively long processing time.

The Constraint-Based Model thus predicts that at the ambiguous pronoun, processing will be intensified relative to unambiguous sentences when the biases are approximately equal, leading to longer processing times. At any later point of disambiguation, there may again be intensified processing, as there will be

some percentage of incorrect initial analyses (predictions summarized in Table 2.1).

Several ERP studies have presented evidence that is in agreement with this prediction, even though this was not the goal of the research. Van Berkum et al. (2007) report that there was a greater NRef effect for ambiguous referents than unambiguous referents, which starts after the point of ambiguity. Nieuwland et al. (2007) found greater activation in a functional Magnetic Resonance Imaging (fMRI) study.

2.1.2.2 *The Unrestricted Race Model*

Some research into syntactically ambiguous sentences has suggested that under the right conditions, sentences with syntactic ambiguities can be *faster* to process than unambiguous sentences (Pickering and Traxler, 1998; Pickering et al., 2000; Van Gompel et al., 2000, 2001; Traxler et al., 2000; Van Gompel et al., 2005). Specifically, globally ambiguous sentences, where the ambiguity is not resolved, appear to be easier to process than sentences that are disambiguated.

Van Gompel et al. (2000) have suggested the *Unrestricted Race Model* to account for cases where globally ambiguous sentences are easier to process. They propose a serial model with two distinct phases. In the first phase, all available information from the context plus the syntactic category information of the current (ambiguous) word is used to construct all possible structures. These structures enter into a race: the structure that gets built first wins the race and is adopted. Immediately after that, the second phase begins in which other (e.g. semantic, pragmatic) information associated with the current ambiguous word can confirm the choice from the first phase, or cause reanalysis.

In globally ambiguous sentences, it does not matter which choice is made in the first phase of the decision process: the sentence remains ambiguous, and no reanalysis is necessary. However, if the sentence is disambiguated at some point, this will cause reanalysis in those cases where the wrong choice was made (which can be about 50% of the time in case of equibias). Under these assumptions, ambiguous structures can be processed faster than disambiguated structures.

In the present paper we will investigate whether the Unrestricted Race Model is a possible model of pronoun resolution. Suppose that in example (1), when the processor reaches the pronoun, there is no strong bias toward either of the characters. The parser will then choose either *Frank* or *Steve* as a referent, and continue forward. It will only slow down if later on in the sentence there is information going against that initial choice. For example if *Steve* is chosen as the referent for (1), this choice will clash with the conceptual information signalled by the word *smile*, since it is unlikely that *Steve* would smile as he was the one who lost at cards. When this clash occurs it will lead to longer processing times as the processor tries interpreting the sentence a second time, now with *Frank* as the referent of the pronoun.

If the reference can be disambiguated on the basis of gender, as in (2), and there is no strong bias to either of the potential referents, it will lead to longer processing times. This is because with no strong bias the coreference will be assigned to *Frank* or *Sarah*. Then, in the second phase, the gender of the pronoun might be taken into account. Since in an equibiased situation *Sarah* will be chosen randomly 50% of the time, if gender is taken into account it would lead to longer processing times on the pronoun, when the pronoun disambiguates gender (2) relative to sentences where the gender does not disambiguate, since in (1) the coreference could never be proven incorrect by gender marking. Thus, the Unrestricted Race Model predicts that referential ambiguity, just as syntactic ambiguity, will lead to faster processing times as compared to disambiguated controls. If subsequent information in the sentence contradicts the earlier choice, this will lead to an increase in processing difficulty.

2.1.3 *Deep and Shallow Processing*

Under both models discussed above, the processor makes a choice in the ambiguous region. However, it is not clear that ambiguities are always completely resolved at all. Ferreira, Bailey, and Ferreira et al., (2002, see also Swets et al., 2008; Ferreira and Patson, 2007; Christianson et al., 2001) have proposed the Good-Enough Hypothesis, drawing on research that sug-

gests that participants do not always process a sentence in full, but rather only as much as is necessary to meet task demands. There has been considerable research demonstrating that some kind of underspecification can occur in processing (e.g. Barton and Sanford, 1993; Sanford and Sturt, 2002). For instance, Barton and Sanford (1993) asked participants to read about an ethical problem and argue about a possible solution. After reading a paragraph introducing the details of an air crash, participants were asked, *After an air crash, where should the survivors be buried?* Many participants in the study did not notice the use of the word *survivor*, which implies *alive*, and even made ethical–legal argumentation based on the sentence. Hence, even when processing a sentence in order to respond to ethical-legal questions, participants did not use information from each individual lexical item in the sentence.

The Good-Enough Hypothesis suggests that manipulation of task could lead to widely varying results, so that for instance reading time patterns on the same set of sentences could appear to favour one model under one task condition, and another model under another. Van Gompel et al. (2001) found evidence that syntactically ambiguous sentences can be easier to process than disambiguated sentences, but Swets et al. (2008) suggest this is because participants did not really attempt to understand the sentence. Van Gompel et al. used comprehension questions; according to Swets et al., this motivated participants to monitor for and remember the verbs, rather than fully processing the sentence. Under the Good-Enough Hypothesis, lack of motivation to process a sentence in full may lead the processor to "withhold judgment" until a later stage of parsing, a process referred to as "buy now, pay later", or may even lead to failure to complete the parse at all. Swets et al. suggest that something like this may have occurred in the experiments reported by van Gompel and colleagues.

2.1.4 *Interim Summary of Model Predictions*

The Constraint-Based Model and the Unrestricted Race Model make specific predictions for the examples presented above (Table 2.1). The Constraint-Based Model predicts longer reaction

times for ambiguous sentences relative to disambiguated controls, both in the ambiguous region and in the region where disambiguating information is encountered. The Unrestricted Race Model, however, predicts that processing in the ambiguous region will be *faster*, not slower than unambiguous control sentences. Both models are in agreement that a later disambiguation will cause higher processing times.

The Good-Enough Hypothesis does not predict any effects specific to ambiguity, but it does suggests that task instructions can determine the specific level of processing in each task, that more intensive tasks will lead to longer processing times, and that a lack of task will lead to shorter processing times.

2.1.5 *Individual Differences*

It has been suggested that individual differences between participants may also influence language processing (Arnold, 1998, 2010; Gernsbacher, 1997). Here, we consider two factors in detail: Reading Span and executive functioning. Several studies have suggested that *working memory* is important in language comprehension (Daneman and Carpenter, 1980; Van den Noort et al., 2008; Waters and Caplan, 1996). Working memory could be integral to pronoun comprehension: what cannot be remembered cannot be available as an antecedent (cf. Van Rij et al., 2013).

We also measured how individuals differed in terms of executive function. Executive function is an umbrella term for the abilities to "control, modify and direct" behaviour (Gioia et al., 2002). The motivation to examine individual differences in executive function is twofold. For one thing, executive functioning is associated with the ability to concentrate on a task, and has been shown to be correlated with the ability to suppress task-unrelated thought (a form of inattentiveness) in research on discourse comprehension (Smallwood and Schooler, 2006; Smallwood, 2007; Smallwood et al., 2008). Secondly, executive functioning can also be linked to the ability to select information from a larger array (e.g. possible referents). We measured working memory using the Dutch version of the reading span task (Van den Noort et al., 2008), and measured executive func-

Table 2.1: Summary of predictions for pronominal processing

MODEL	REGION	
MODEL	PRONOUN AND SPILLOVER	DISAMBIGUATING PHRASE AND SPILLOVER
Constraint-Based	Amb > Unamb	Amb > Unam
Unrestricted Race*	Amb < Unamb	Amb > Unamb
Good-enough	Disamb. ≠ Plaus. ≠ SPR-only	

Note: Ambiguous pronoun = Amb, Unambiguous pronoun = Unamb Disambiguation Question = Disamb, Plausibility Rating=Plaus. *If gender morphology is considered part of syntactic information.

tion using a Dutch version of the Behavioural Rating Inventory of Executive Function–Adult self-report (BRIEF–A) (Roth et al., 2005).

2.2 THE PRESENT STUDY

In order to test how models of sentence processing apply to pronoun comprehension, we constructed short stories, equivalent to (Table 2.2) in form. The critical difference between the models is how they handle a point of ambiguity where the bias toward the referents is very nearly equal. For this reason, it was crucial to create a point of rough equibias at the pronoun. This was done by using referents that had both been subjects (parallelism of syntactic roles), one of which was referred to first (first mentioned bias) and one of which was the more recent subject (recency or current topic). In that circumstance, the Unrestricted Race Model predicts that an ambiguous pronoun would be faster. At a point of disambiguation, the previously ambiguous condition is predicted to be slower, as an initially chosen antecedent will be proven wrong in approximately half of the cases. In contrast, the Constraint-Based Model predicts that there will be longer reaction times RTs at points of roughly equal plausibility, but also slowing at places of disambiguation.

Table 2.2: Example of stimulus materials.

AMBIGUOUS PRONOUN	UNAMBIGUOUS PRONOUN
Frank played cards with Steve for a huge sum. Steve lost to Frank.	*Frank played cards with Sarah for a huge sum. Sarah lost to Frank.*
CRITICAL SENTENCE	
/He / had a big / smile / when leaving.	

Note. Pronoun disambiguation region and pragmatic disambiguation region indicated with slashes (i.e., / *smile* / or / *frown* /).

2.2.1 *Methods*

2.2.1.1 *Participants*

Seventy-seven participants were recruited from the University of Groningen community (mean age 22, 18–32, 19 male). Participants received 5 euros as compensation. None had neurological impairments.

2.2.1.2 *Materials*

The experimental materials were comprised of 32 three-sentence stories, interspersed with 24 distractor items (from a different experiment). Table 2.2 illustrates the experimental conditions for one item. Materials can be found in Appendix A.

AMBIGUITY. The gender of the characters in the story was varied in order to manipulate initial ambiguity (see Table 2.2). Either both characters were male, making the pronoun ambiguous, or there was one male and one female character allowing the referent to be identified from the pronoun alone. In addition to Ambiguity, we counterbalanced for order-of-mention. Order-of-Mention was added so that the pronoun would not always predictably refer to First Mention (e.g. *Frank*) or Second Mention (e.g. *Steve*) within the experiment, but would be unpredictable. Following the research of Gernsbacher et al. (1989), the activation of *Frank* and *Steve* is predicted to be roughly equal at the pronoun in the third sentence, since *Frank* was first mentioned, but *Steve* is the subject of the most recent clause. We manipu-

lated Ambiguity (Ambiguous, Unambiguous), as shown in Table 2.2.

MATERIALS PRETESTS. We carefully pretested the plausibility of the preceding context and of the disambiguating regions. These pretests were important because the differences between the three models rely crucially on approximate equiplausibility at the pronoun, followed by a phrase that effectively disambiguates the pronoun. The conceptual plausibility of the referents was measured with a test of the plausibility of each potential antecedent at the pronoun and another test with the disambiguating phrase.

PRETEST 1. We measured the continuation preference of First Mention and Second Mention as a continuation of the last sentence. This was accomplished by replacing the last sentence as shown in (4).

(4) *Frank played cards with Steve for a huge sum.*
Steve lost all his money to Frank.
Frank...<did something>

Participants were asked to rate the character for preference as a continuation of the story using a seven-point Likert-type scale. One was "very improbable" and seven was "very probable" (since probable is the closest equivalent to "plausible" in Dutch). A given participant only saw one character as a possible continuation. This was intended to measure the accessibility of the two possible antecedents at this point in the sentence. There was a preference for First Mention (e.g. *Frank*), as continuation of the story (5.8, 0.5) although Second Mention was also considered acceptable (e.g. *Steve*), (3.9, 0.7). This preference was significant ($p < .001$), probably partly because repeating the full name *Steve* to refer to the second mention character was not considered very felicitous (Results by item group in Table 2.3).

PRETEST 2. It was also necessary to ensure that all of the disambiguating phrases did disambiguate to the intended referent, and to that referent exclusively. In Pretest 2, we measured how plausible each scenario was for First Mention (5c & 5d) and Sec-

Table 2.3: Matching of stimuli by item group based on pretest 1.

ITEM GROUP	FIRST MENTION	SECOND MENTION
1	5.62 (0.70)	3.96 (0.75)
2	5.80 (0.50)	3.91 (0.89)
3	5.87 (0.54)	3.59 (0.42)
4	5.99 (0.48)	4.06 (0.86)

Note. Mean (SD).

ond Mention (5e & 5f). Participants were explicitly instructed to rate the plausibility of the event, rather than whether the sentence was a good continuation of the discourse.

(5) a. *Frank played cards with Steve for a huge sum.*
b. *Steve lost all of his money to Frank.*
c. *Frank had a big smile when leaving.*
First Mention- Correct character 5.64 (.77)
d. *#Frank had a big frown when leaving.*
First Mention- Incorrect character 2.10 (.89)
e. *Steve had a big frown when leaving.*
Second Mention-Correct character 5.54 (.99)
f. *#Steve had a big smile when leaving.*
Second Mention-Incorrect character 2.36 (.74)

The plausibility scores for the items included in the final experiment are shown in (5). The First Mention and Second Mention disambiguation manipulations were successful, since 5c is more plausible than 5d, and 5e is more plausible than 5f (results by item group in Table 2.4).[1]

SUMMARY OF ITEM MATCHING. Four stimuli groups with eight items each were created for distribution across lists. These groups were matched in terms of Pretest 1, Pretest 2 (Table 2.4), the frequency of the character names, established on the basis of word frequency (Table 2.5), and the distance from the pro-

1 We discovered that 11 items did not correctly disambiguate the referent and so made 11 new items and re-piloted these. Of these 11, four met our requirements for inclusion, namely that the preference for the correct referent must be above 5, for the incorrect referent below 5, with a difference of at least 1.

Table 2.4: Matching of stimuli by item group based on pretest 2: Plausibility of character as referent of disambiguating phrase

PHRASE BIAS TOWARD FIRST MENTION (FRANK)			
ITEM GROUP	CORRECT NAME (*Frank*)	INCORRECT NAME (*Steve*)	DIFFERENCE (Corr-Incorr)
1	5.9 (0.6)	2.3 (0.6)	3.6 (1.0)
2	5.5 (0.7)	2.0 (0.9)	3.5 (0.6)
3	5.7 (0.9)	2.1 (1.2)	3.6 (0.9)
4	5.5 (1.0)	2.0 (0.9)	3.5 (1.0)
PHRASE BIAS TOWARD SECOND MENTION (STEVE)			
ITEM GROUP	CORRECT NAME (*Steve*)	INCORRECT NAME (*Frank*)	DIFFERENCE (Corr-Incorr)
1	5.7 (0.6)	2.4 (0.6)	3.3 (1.1)
2	5.6 (1.0)	2.4 (1.1)	3.2 (1.0)
3	5.3 (1.0)	2.3 (0.4)	3.3 (1.1)
4	5.3 (1.4)	2.3 (0.9)	3.0 (1.3)

Note. Mean (sd) of stimuli.

noun to the last occurrence of each name the length and frequency of the target verb and following words were counterbalanced across conditions. The differences between the item groups was then tested with an ANOVA, taking item group as a between items variable and Order-of-Mention as a within items variable. The item groups were matched according to pretest 1: both the plausibility of NP1 and NP2 were considered in the matching (Table 2.3: $p = .34$ and $p = .48$ respectively).

As mentioned in the discussion of pretest 2, the item groups were matched on the basis of the difference between the plausibility of the references (Table 2.4, all p's $> .43$), as well as on the plausibility of the references themselves (Table 2.4, all p's $> .42$;). The difference in plausibility between the referents has been shown to have an effect on speed of pronoun comprehension (Hirst and Brill, 1980), so we wanted all groups to have an equal amount of referent bias. The average referent bias in our experiment was comparable to the *strong bias* category in Hirst and Brill (1980; biases of 2.84 and 3.80 in their experiment).

Table 2.5: Matching of stimuli by item group based on log Google name frequencies.

ITEM			DIFFERENCE	
GROUP	MAN 1	MAN 2	1 — 2	WOMAN
1	6.14 (0.48)	5.90 (0.49)	0.23 (0.25)	5.87 (0.38)
2	6.12 (0.20)	6.00 (0.25)	0.12 (0.26)	5.98 (0.27)
3	6.15 (0.53)	6.01 (0.32)	0.14 (0.34)	5.96 (0.49)
4	6.14 (0.30)	6.05 (0.25)	0.08 (0.42)	5.96 (0.21)

Note. Mean (sd). In Table 2.2, Man 1 = *Frank*, Man 2 = *Steve* and Woman = *Sarah*

Included names were all common in the north of the Netherlands, with a Google frequency of 64,000-plus. The names were matched in terms of frequency and on the difference in frequency between the names, to balance any effect of proper name frequency across conditions (Table 2.5, all p's > .81). The distance between the First Mention name in the last sentence and the pronoun was also matched, in order to control the amount of interference caused by Character name 2 (p = .905).

2.2.1.3 *Procedure*

Participants arrived at the laboratory and were given instructions before starting the experiment. Participants initially saw the three sentences with all letters replaced by dashes. Pushing the space bar revealed the text one word at a time. With each press of the button, a subsequent word would immediately change from dashes to letters, and the previous word would immediately revert to dashes. After the reading experiment, participants were administered the reading span task and completed the executive function inventory.

TASK MANIPULATION. In Disambiguating Question task, an experimental story such as for example (6) would be followed by question such as (6a), in the Plausibility Rating task a plausibility rating such as (6b); in the SPR-Only condition there were no questions following the stimuli.

(6) *Frank played cards with Sarah for a huge sum.*
Sarah lost all her money to Frank.
He had a big smile when leaving the room.

a. Disambiguating Question:
Who had a big smile when leaving the room? (Sarah/Frank)
b. Plausibility Rating:
Very implausible *1* *2* *3* *4* *5* *Very plausible*

Following completion of the self-paced reading portion of the study, the participants were debriefed. Particular attention was paid to asking the participant what they thought the experiment was about. The participant was then asked to complete a survey of executive function abilities, the BRIEF–A.

BRIEF–A. Executive function was measured with the Behavioural Rating Inventory of Executive Function–Adult self-report (Roth, Isquith, Gioia, 2005; 2007). The BRIEF–A has participants answer questions about their executive function abilities in day–to–day life. It measures both behavioural regulation and metacognition. Participants in the three task conditions did not differ significantly with respect to BRIEF–A scores (One-way ANOVA, $p = .29$).

READING SPAN. Participants were administered the reading span task (van den Noort, et al., 2008) in the way Friedman and Miyake prescribe to increase the efficacy of the test (2004; 2005). Participants in the different conditions were matched in terms of reading span across task conditions (One-way ANOVA, $p = .67$)

2.2.2 *Results*

One item was discarded due to a recording error. The response times were prepared for analysis by discarding responses under 40 ms and over 4000 ms, with retention of 98.7% of the data. We then capped the data at 2.5 standard deviations from item means or subject means, so that outliers would not have an overly strong effect on the results. The mean reading times appear in Figures 2.1–2.4. We performed by-subject and by-

item repeated measures ANOVAs at the critical words, and we included item group, list, and order-of-mention in the analysis in order to reduce variance. We examined the pronoun, one word after the pronoun, one word before the disambiguating region, the disambiguating region, and one word after. The *regions* comprised one word. The highest-level interactions are presented in Table 2.6, with analyses by region in Table 2.7. Because of the significant interaction of Task, Position, and Ambiguity, we were licensed to examine the effects by task and by position (Table 2.8). We will present the results of the statistical analyses per region.

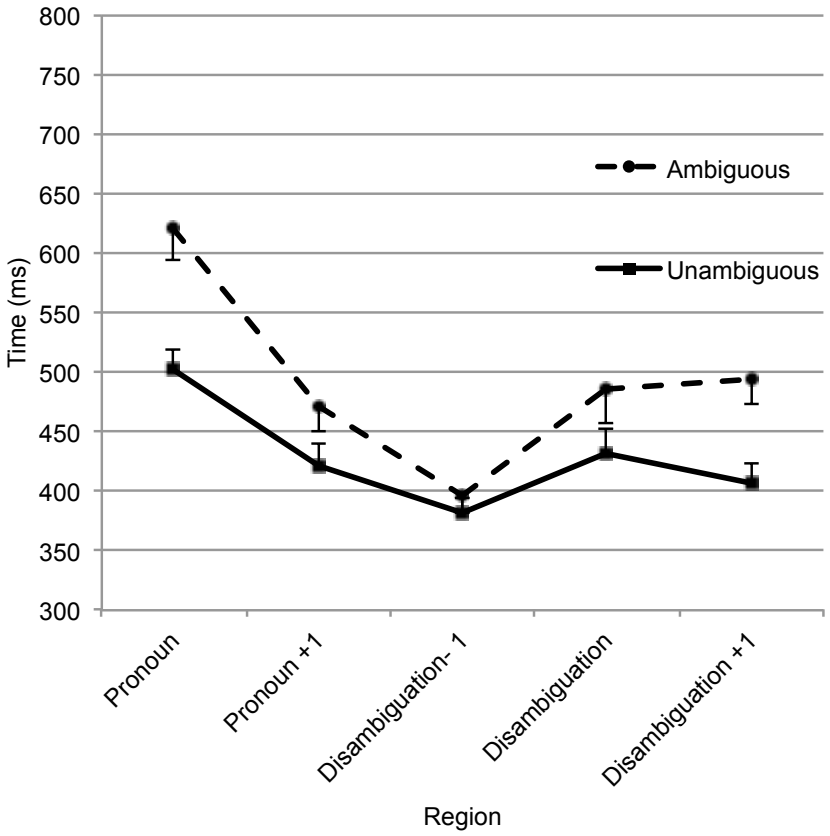

Figure 2.1: Mean reaction times and standard errors for the Disambiguation Question task.

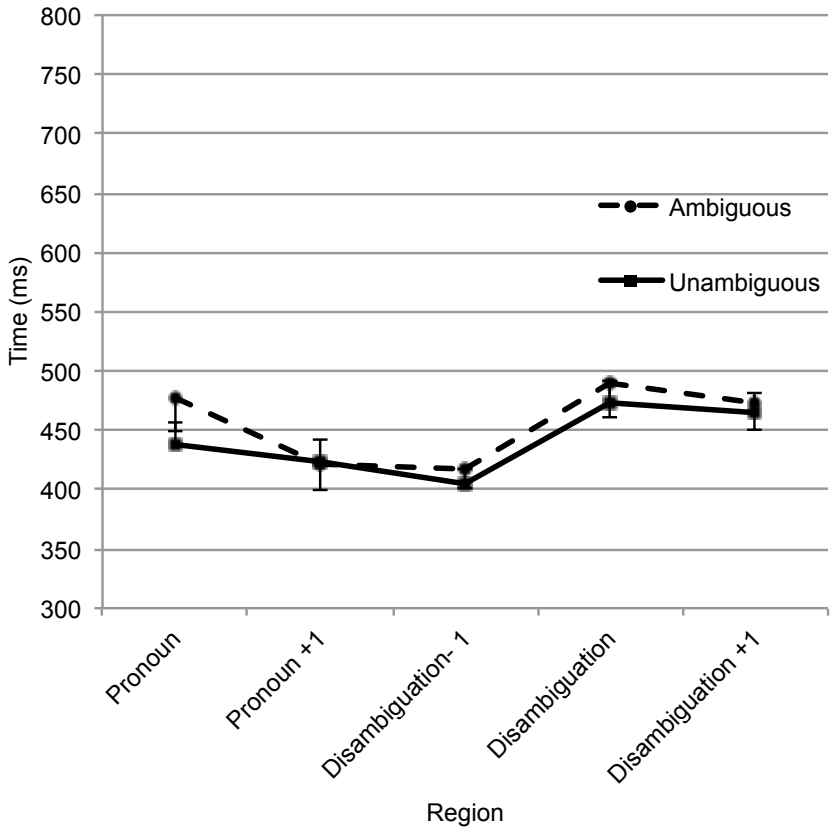

Figure 2.2: Mean reaction times and standard errors for the Plausibility Rating task.

2.2.2.1 *Ambiguous Region (pronoun plus one extra word)*

The pronoun region and one word after the pronoun showed significant main effects of ambiguity (Figures 2.1–2.3). These main effects were qualified by an interaction of Ambiguity and Task condition. Statistical inference indicates that this is attributable to the absence of an effect in the SPR-Only experiment (Figure 2.3), but significant effects of Ambiguity in the Disambiguation Question and Plausibility Rating tasks (Figures 2.1 and 2.2). In both the Disambiguation Question and the Plausibility Rating question levels, this was due to Ambiguous taking longer to process than Unambiguous.

Follow-up analyses of Task by position (Table 2.9, Figure 2.4) show that at the pronoun (but not at the pronoun +1 region), there is a pattern of Disambiguation Question > Plausibility Rating > SPR–only.

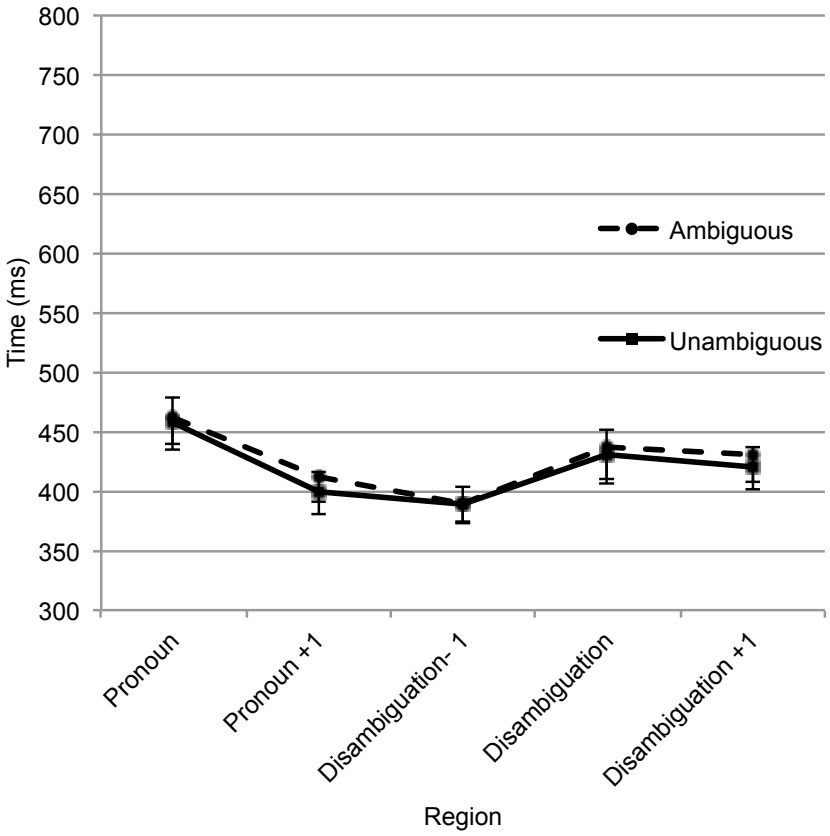

Figure 2.3: Mean reaction times and standard errors for the SPR-Only task.

2.2.2.2 *Disambiguation Region*

There was a main effect of Ambiguity starting in the Disambiguation region and continuing into the spillover region (Figure 2.1). In the spillover this main effect was qualified by an interaction of Ambiguity and Task. These effects appear to result from an ambiguity disadvantage in the Disambiguation task, but not in the other two tasks (Figures 2.2 and 2.3). This may be driven by the ambiguity disadvantage continuing into the spillover region in the Disambiguating Question condition, but not in the other task conditions.

Turning to the follow-up analyses of the task by position interaction (Figure 2.2.4, Table 2.9), in Disambiguation region -1, Plausibility Rating task RTs were significantly greater than in the Disambiguation Question task and SPR-Only RTs. In the Disambiguation region itself, RTs in the Plausibility Rating task

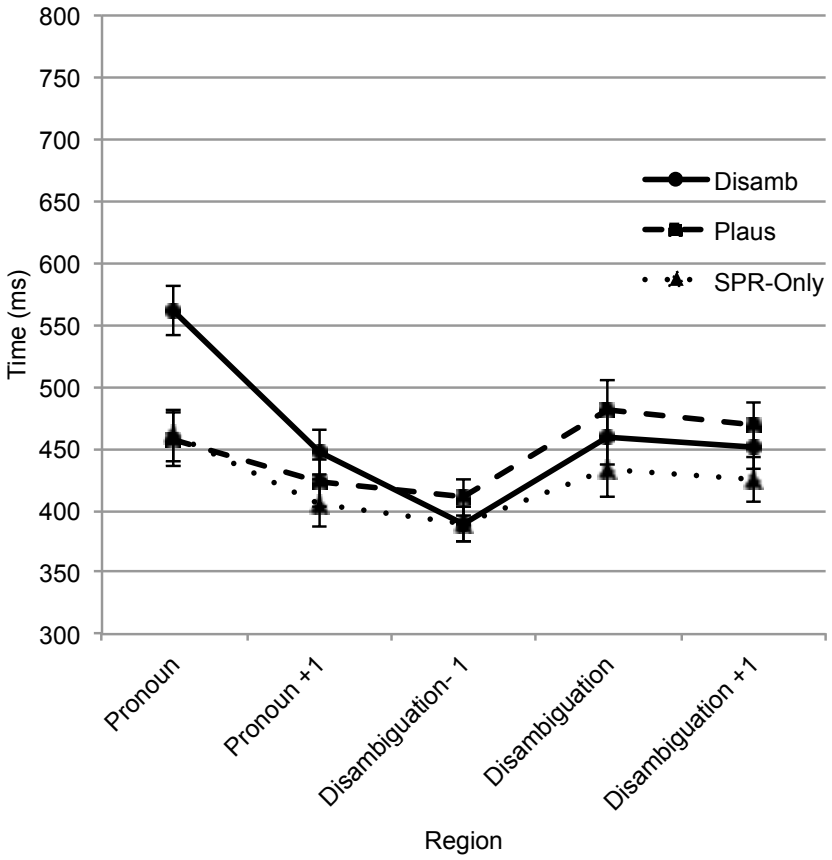

Figure 2.4: Mean reaction times and standard errors by Task; Disambiguation Question = Disamb, Plausibility Rating=Plaus. Ambiguous and Unambiguous conditions are combined.

were greater than in SPR-Only. There was no effect of Task in the Disambiguation +1 region.

2.2.2.3 *Individual Differences*

In order to examine the interaction between individual differences and ambiguity resolution, we performed correlations between the RT-difference between ambiguous and unambiguous stimuli (in the Pronoun region and the Disambiguation region) and the measures of working memory (reading span) and executive function (BRIEF–A). There were no significant correlations with reading span or BRIEF–A scores. [2]

2 No measure was significant for both the Pronoun and the Disambiguation regions, or even both NP1 referring pronouns and NP2 referring pronouns.

Table 2.6: Task–level Interactions.

	df	F	p	η^2
Task	$F_1(2, 65)$	1.35	.268	.04
	$F_2(2, 54)$	14.25	.000	.35
Region	$F_1(4, 260)$	37.65	.000	.37
	$F_2(4, 108)$	12.65	.000	.32
Ambiguity	$F_1(1, 65)$	25.90	.000	.29
	$F_2(1, 27)$	33.44	.000	.55
Task Region	$F_1(8, 260)$	7.96	.000	.20
	$F_2(8, 216)$	8.86	.000	.27
Task x Ambiguity	$F_1(2, 65)$	10.37	.000	.24
	$F_2(2, 54)$	17.05	.000	.39
Region x Ambiguity	$F_1(4, 260)$	3.47	.009	.05
	$F_2(4, 108)$	4.72	.002	.15
Task x Region x Ambiguity	$F_1(8, 260)$	2.10	.035	.06
	$F_2(8, 216)$	2.06	.041	.07

2.3 GENERAL DISCUSSION

2.3.1 *Serial and Parallel Models*

The goal of this study is to compare how well Serial (Unrestricted Race Model) and Parallel (Constraint-Based Model) parsers predict reading times of sentences with pronominal ambiguity. The Constraint-Based Model suggests that ambiguity will cause increased RTs as compared to unambiguous pronouns, whereas the Unrestricted Race model predicts the opposite effect. Both models predict longer reading times at a later point of disambiguation (Table 2.10). We additionally manipulated task, to find out whether the different results for ambiguity found in the literature may have been due in part to task effects.

We found longer RTs for ambiguous pronouns (with two equiplausible referents) in the two tasks that required (relatively)

The only correlations were between NP1 referring pronouns in the disambiguation region and a sub-measure of the BRIEF–A that measures planning and organization ability (r = .25, p<.05) and NP2 referring pronouns in the pronoun region, and a sub-measure of the BRIEF–A that measures inhibition (r = -.28, p<.005).

Table 2.7: Analysis of Variance for Interactions.

		WORD REGION				
		PRONOUN	+1	-1	PRAG. DISAMB.	+1
Ambiguity	$F_1(1, 65)$	18.66	8.02	2.87	4.5	11.97
	p	.000	.006	.095	.037	.001
	η^2	.22	.11	.042	.065	.155
	$F_2(1, 27)$	25.41	11.28	4.23	8.57	10.68
	p	.000	.002	.049	.007	.003
	η^2	.485	.30	.14	.24	.28
Task	$F_1(2, 65)$	8.48	1.24	.76	1.04	1.42
	p	.001	.296	.471	.360	.249
	η^2	.21	.04	.02	.03	.04
	$F_2(2, 54)$	58.00	17.28	8.11	8.43	1.65
	p	.000	.000	.001	.001	.20
	η^2	.682	.39	.231	.24	.058
Task x	$F_1(2, 65)$	7.70	4.62	.63	1.46	6.6
Ambiguity	p	.001	.013	.019	.241	.002
	η^2	.19	.13	.54	.24	.17
	$F_2(2, 54)$	12.21	4.92	1.29	1.70	5.41
	p	.000	.011	.284	.194	.007
	η^2	.31	.15	.05	.06	.17

Note. Prag. Disamb = Pragmatic Disambiguation.

Table 2.8: Analysis of Variance by Task Manipulation.

DISAMBIGUATION QUESTION						
		WORD REGION				
		PRONOUN	+1	-1	PRAG. DISAMB.	+1
Ambiguity	$F_1(1, 24)$	14.74	8.79	3.63	3.71	15.47
	p	.001	.007	.07	.06	.001
	η^2	.38	.27	.131	.13	.39
	$F_2(1, 27)$	22.09	10.96	3.47	6.75	11.35
	p	.000	.003	.073	.015	.002
	η^2	.45	.29	.11	.20	.30
PLAUSIBILITY RATING						
		WORD REGION				
		PRONOUN	+1	-1	PRAG. DISAMB.	+1
Ambiguity	$F_1(1, 21)$	12.56	.07	.92	1.11	.20
	p	.002	.007	.349	.303	.657
	η^2	.37	.00	.04	.05	.01
	$F_2(1,27)$	10.92	.04	2.62	.82	.53
	p	.003	.853	.117	.373	.471
	η^2	.29	.00	.09	.03	.02
SPR–ONLY						
		WORD REGION				
		PRONOUN	+1	-1	PRAG. DISAMB.	+1
Ambiguity	$F_1(1, 20)$	.081	2.415	.008	.28	1.62
	p	.779	.136	.929	.606	.218
	η^2	.00	.11	.00	.61	.08
	$F_2(1, 27)$	.03	2.14	.036	.09	.45
	p	.865	.155	.851	.763	.509
	η^2	.00	.07	.00	.00	.02

Note. Prag. Disamb = Pragmatic Disambiguation.

Table 2.9: Bonferroni post-hoc comparisons between Task effects.

		WORD REGION				
		PRONOUN	+1	-1	PRAG. DISAMB.	+1
Disamb. vs. Plaus.	Subject	.002 D>P**	1.00	.822	1.00	1.00
	Items	.000 D>P***	.002 D>P**	.013 P>D*	.149	.157
Disam. vs. SPR-Only	Subject	.003 D>S**	.372	1.00	1.00	.940
	Items	.000 D>S***	.000 D>S***	1.00	.228	1.00
Plaus. vs. SPR-Only	Subject	1.00	1.00	.914	.463	1.00
	Items	1.00	.034 P>S*	.002 P>S**	.000 P>S***	.489

Note: Disambiguation Question = Disamb, Plausibility Rating=Plaus. Prag. Disamb. = Pragmatic Disambiguation

extensive processing, but not in the experiment without a secondary task. This represents clear evidence in favour of the Constraint-Based Model. In the disambiguation region, both the Constraint-Based Model and the Unrestricted Race Model predicted higher processing times in the ambiguous conditions. However, ambiguity effects in the Disambiguation region only occurred in the experiment with the Disambiguating Question task, and not for the other two tasks. We discuss these effects in relation to underspecified processing in the following section.

2.3.2 *The Good-Enough Hypothesis and Task Effects*

The Good-Enough Hypothesis predicts that differences between tasks will lead to differences in how stimuli are processed. There were readily apparent significant differences between the task conditions in terms of processing time on all but the spillover region of the disambiguation region (Pragmatic Disambiguation +1). In the pronoun regions (Pronoun Disambiguation and +1),

Table 2.10: Summary of predictions and results for pronominal processing

PREDICTIONS		
	REGION	
MODEL	PRONOUN AND SPILLOVER	PRAG. DISAMB. AND SPILLOVER
Constraint-Based	Amb > Unamb	Amb > Unam
Unrestricted Race*	Amb < Unamb	Amb > Unamb
Good-enough	Disamb. ≠ Plaus. ≠ SPR-only	
RESULTS		
MODEL	REGION	
TASK	PRONOUN AND SPILLOVER	PRAG. DISAMB. AND SPILLOVER
Disamb	Amb > Unamb	Amb > Unamb
Plaus.	Amb > Unamb	n.s.
SPR-only	n.s.	n.s.
Task effect by region	Disamb > Plaus	Plaus > SPR-Only
	Disamb > SPR-Only	n.s.

Note: Ambiguous pronoun = Amb, Unambiguous pronoun = Unamb Disambiguation Question = Disamb, Plausibility Rating=Plaus. *If gender morphology is considered part of syntactic information. Prag. Disamb. = Pragmatic Disambiguation

the order from longest processing time to shortest was Disambiguating Question > Plausibility Rating > SPR-Only. This suggests that the task manipulation had an effect on the effort that our participants were putting into processing the sentences (Table 2.9).

The relationship between these task effects changed throughout the pragmatic disambiguation regions. In Pragmatic Disambiguation -1, the Plausibility Rating task actually had longer RTs than the Disambiguation Question task and SPR-Only. This suggests that the Plausibility Rating task caused more spillover processing following the Pronoun Region than the other two task conditions.

At the Pragmatic Disambiguation itself, the Plausibility Rating task showed statistically longer RTs than SPR-Only, and was statistically indistinguishable from Disambiguating Question—even though there was no evidence of an ambiguity disadvantage for Pragmatic disambiguation in this region.

The results broadly endorse the predictions of the Good-Enough Hypothesis that task requirements shape processing. They also refute the notion that some Task conditions may yield results that look like the Unrestricted Race Model. In fact, none of the models precisely foresaw the Task effects that materialized.

In the Plausibility Rating task it appears that participants initially paid attention to the pronoun interpretation, but then paid greater attention at the disambiguation region, since there was an ambiguity effect in the pronoun region, and in the pragmatic disambiguation region response times were high in both ambiguous and unambiguous conditions. Moreover, the average RTs for the region were higher than SPR-Only, and indistinguishable from the Disambiguation Question task. The results from the Plausibility Rating task suggest that shallower processing in the pronoun regions was compensated for with deeper processing in the disambiguation regions, perhaps as participants mentally double checked hasty referent choice. The Plausibility Rating task thus provides marginal evidence that somewhat supports the "buy now, pay later" idea suggested by the Good-Enough Hypothesis. In contrast to the Plausibility Rating task, the SPR-Only task showed no evidence of greater processing at a later time, which is consistent with the lack of response.

However, the Good-Enough Hypothesis did not predict task requirements with a high degree of precision. For example, why was there no ambiguity disadvantage in the Pragmatic Disambiguation Region? We argue that the ambiguity disadvantage was absent in these regions due to ceiling effects. In both the Disambiguating Question condition, and the Plausibility Rating task, we found significant ambiguity disadvantages at the pronoun. In the pragmatic disambiguation region, there was an arithmetic ambiguity disadvantage that did not reach significance.

One interpretation is that in the Plausibility Rating task, the parser was using the pragmatic disambiguation, irrespective of whether the pronoun disambiguated the sentence or not. There is further evidence for this suggestion, because at the pronoun in the Plausibility Rating task the effect of ambiguity was less significant, with a lower effect size. This suggests that the parser did not always take into account the initial pronoun disambiguation, perhaps since the information was orthographically minimal—the words for *he* and *she* in Dutch are *hij* and *zij*, a difference of one letter, which is not much if you are in a rush, and very different from the > 4 letters of difference that was present in the plausibility disambiguation region. Alternately, Garnham et al. (1992) suggested that participants use gender strategically, sometimes weighting it higher than others.

2.3.3 *Individual Differences*

The individual differences that we examined only reached significance in sub-tests of the executive function inventory, and were not interpretable, in contrast with past studies on children. These null results could be taken as evidence that language is ergonomically designed for adults: Individuals that had working memory and executive function within normal bounds did not show reliable effects of differences between them. However, all of our participants were university students within the normal bounds of executive function and working memory. It seems possible that individuals with lower working memory or executive function ability may have greater difficulty with ambiguity. Examination of populations with lower working memory

ability, such as Alzheimer's patients could reveal very different results—perhaps yielding more extreme differences between unambiguous and ambiguous, or complete failure to retrieve a referent.

2.4 CONCLUSIONS

This experiment found support for the Constraint-Based Model in pronoun ambiguity resolution. We observed that when participants were closely attending to the stimuli, there was an ambiguity disadvantage for pronominal processing. However, when task demands are less stringent, participants appear to become less focussed and process on a more superficial level.

This study adds to past research by presenting further evidence that the processor has different modes, typified by the allotment of more processing time to regions that help satisfy task demands. Several different processing modes can be characterized based on the pattern of results. It appears that the Constraint-Based Model best describes attentive, natural processing, and that otherwise participants engage in processing that is only good-enough.

The evidence that we found corroborating the Constraint-Based Model is part of a broader trend. The past several years have seen increasing dominance of parallel processing models in language processing, a trend which appears to be a shift away from the serial processing models of the past (Spivey, 2008). Evidence for competition has most notably come from the visual world paradigm, in experiments looking at lexical and syntactic ambiguities. However, support for competition from self-paced reading and reading studies of eye-tracking has been thin on the ground, or has found evidence that contradicts parallel processing outright (see Hoeks and Hendriks, 2011 for additional criticism).

However, Spivey (2008) contends that competition effects have not been found in reading tasks since reaction times do not show decision processes as they unfold, as do, for example visual world studies and ERP studies. A corollary to this argument is that decades of psycholinguistic research should be pitched.

This present study, in contrast, suggests that there are a few babies that can be saved from the SPR bathwater.

In particular, our results suggest that discrepancies between SPR and other methodologies (and from one SPR study to another) can be explained by task effects, since SPR and visual world tend to vary substantially in terms of task involvement. The difference in degree of task engagement between visual world studies and the bulk of SPR studies could hardly be greater. In the original visual world study rather than passively clicking a spacebar, participants followed instructions that required them to perform a visual search, reach out, grasp a toy, and move it into place (Tanenhaus et al., 1995). Lack of engagement thus seems to be an important difference between visual world studies and most SPR studies.

Thus, deep processing of linguistic stimuli may be a requirement for competition effects to occur. The reason for this may be that deep and superficial processing require (slightly) different functioning of the human sentence processor. Alternately, what visual world tasks and pronoun resolution have in common is that they require the alternatives to become more or less conscious for the language processor. Maybe this feature is decisive for competition effects to occur. Future research is necessary to investigate these explanations further. Our results imply that it is necessary to understand what is happening not only in high-engagement language comprehension, but also in low-engagement language comprehension, and that it makes sense to look at the specific requirements of different tasks. In any case, we have shown that competition effects exist in pronoun resolution.

3

COMPREHENSION OF MARKED PRONOUNS IN SPANISH AND ENGLISH

Interaction of World Knowledge and Markedness in Object anaphor interpretation

ABSTRACT

Previous research on pronoun resolution has identified several individual factors that are deemed to be important for resolving reference. In this paper we will argue that of these factors, as tested here, *plausibility* is the most important, but interacts with *form markedness* and *structural parallelism*. We investigated how listeners resolved *object pronouns* that were ambiguous in the sense of having more than one possible antecedent. We manipulated the *form* of the anaphoric expression in terms of accentuation (English: Experiments 1a and 2a) and morphology (Spanish: Experiments 1b and 2b). We looked at sentences where both antecedents were equally plausible, or where only one of the antecedents was plausible. Listeners generally resolved toward the (parallel) grammatical object of the previous clause. When the pronouns were marked due to accentuation (English) or use of specific morphology (Spanish), preference switched to the alternative antecedent, the grammatical subject of the previous clause. In contrast, when one of the two antecedents was a much more plausible antecedent than the other, antecedent choice was almost wholly dictated by plausibility, although reference form prominence did significantly attenuate the strength of the preference.[1]

1 This chapter is a modified version of Taylor et al. (2013)

3.1 INTRODUCTION

Ambiguous pronouns abound in language use, presenting a challenge to the human language processor. Multiple factors are assumed to be involved in pronoun resolution (see Poesio et al., 2011; Arnold, 2010 for reviews of relevant factors). The goal of the present paper is to examine three factors that we deem highly representative: How emphatic the referring expression is, the structure in which it appears, and the plausibility of the available referents. There have been several claims about which of these factors might be the most important, although less research simultaneously investigating all three.

We examine form markedness through cross-linguistic comparison of English and Spanish; to see whether pitch accenting, or morphological markedness can change who the anaphor refers to when both antecedents are plausible. In particular, we look at whether there is a preference for parallel interpretation, whereby object pronouns would preferably be resolved to the grammatical object of the previous clause (Smyth, 1994; Kehler et al., 2008, e.g.), or a subject bias, whereby the pronoun would refer to the subject of the previous clause (Crawley et al., 1990). We then look at what happens when there is a bias toward the subject or object, to see how structure and form markedness interact with bias. Before discussing our experiments in detail, we present an overview of the psycholinguistic research on antecedent markedness, grammatical roles and plausibility in pronoun comprehension.

3.1.1 *Reference Form Markedness and Antecedent Selection*

In general, when a referent is harder to identify, it is discussed using a more elaborate form. From the standpoint of the hearer, forms that are longer, or less frequent than expected, will cause hearers to reject the normally preferred antecedent, and search for an alternative one within the previously mentioned entities in discourse (Levinson, 1991; Huang, 1991, 2000; Blackwell, 2001). This appears to be the case in example (2) where accenting switches a preference for *Frank* (as assumed in example 1) to

a preference for *John*:

(1) *John called Frank and Steve emailed him.* (*him* = *Frank*)

(2) *John called Frank and Steve emailed* HIM. (HIM = *John*)

There have been several studies suggesting that when hearers encounter accented pronouns, they choose an entity that is dispreferred in the discourse, relative to unaccented pronouns (1971; see experiments on subject pronouns by Hirschberg and Ward, 1991; Venditti et al., 2001 and Tavano and Kaiser, 2008). While this issue was first brought to light by Lakoff in the 60's, empirical research into accented pronouns is relatively recent. For example, Venditti et al. (2001) studied the difference between accented pronouns and unaccented pronouns by looking at ambiguous subject pronouns, where the pronoun is congruent with either of two previous antecedents; in (3), either *zebra* or *pig* can be the antecedent of *he*. They found that in the unaccented condition (3-b), listeners preferred the pronoun to refer to the subject (the zebra). In the accented condition (3-c), however, the prepositional object (the *pig*) was the preferred antecedent.

(3) a. *The zebra put a bucket of soapy water next to the pig near the front of the car.*

b. *Then he got out some sponges.*

c. *Then* HE *got out some sponges.*

This finding suggests that accenting a subject pronoun has the effect of *switching* the preferred antecedent. The present study investigates whether this also holds true for *object* pronouns. Investigating object pronouns is important because, as we will discuss below, the way object pronouns are interpreted allows us to address several other factors which affect antecedent choice, and thus how these interact with reference form.

Languages can differ strongly in the ways in which a referring expression such as a pronoun can be made marked. In English, pronouns can receive a pitch accent (as in 2). However, in languages such as Spanish, accenting pronouns for emphasis seems

to be rather dispreferred, or at least highly infrequent. Instead, *pronoun doubling* is used to provide emphasis. Due to it being a more marked form, an accented pronoun HIM may be taken to refer to an entity that is different than that chosen by the unaccented pronoun him, leading to a change in preference of antecedent and even to an *antecedent switch*.

There is considerable cross-linguistic variation in how emphasis can be achieved, and this produces one of the reasons for focusing on object pronouns. In English, emphasis for both subject pronouns and object pronouns is analogous, irrespective of grammatical role: accenting the pronoun is the major way to make it more marked. However, in Spanish, the situation is quite different: in object position, object clitics like *la*, *lo*, or *le*, must be used, as in (4) below. A null form here would be ungrammatical. As shown, object clitics must be present in transitive sentences with no full noun in the object position. As indicated by the subscript 2, the antecedent is normally taken to be the preceding object, *Mónica*. (In the examples below, a subscript "1" refers to the subject, *Sandra* and subscript "2" refers to the object, *Mónica*).

(4) *Sandra$_1$ llam-ó a Mónica$_2$, y Rogelio* le$_2$ /
Sandra$_1$ call-3sg.PST PREP Mónica$_2$, and Rogelio 3:OBJ$_2$ /
*Ø *escribi-ó.*
*Ø wrote-3sg.PST
'Sandra$_1$ called Mónica$_2$, and Rogelio wrote her$_2$ / *Ø$_2$.'

In the object position in Spanish, emphasis is produced by pronoun doubling, or clitic doubling (shown in example 5). Pronoun doubling, just like accenting in English, appears to have important consequences for pronoun interpretation. (Blackwell, 2001; Baauw et al., 2011) predicts that when a double object pronoun is produced as in (5b), it is a more marked form, and therefore might refer to a different antecedent than in (5a). In (5b), the antecedent *Sandra* should be chosen instead of *Mónica*. (Note that animate referents in Spanish are obligatorily indirect objects when in the object position, taking the preposition *a*, as in (5b).

(5) a. *Sandra$_1$ llam-ó a Mónica$_2$, y Rogelio*
Sandra$_1$ call-3sg.PST PREP Mónica$_2$, and Rogelio
le$_2$ escribi-ó.
3:OBJ$_2$ wrote-3sg.PST
'Sandra$_1$ called Mónica$_2$ and Rogelio wrote her$_2$.'

b. *Sandra$_1$ llam-ó a Mónica$_2$, y Rogelio*
Sandra call-3sg.PST PREP Mónica, and Rogelio
le$_1$ escribi-ó a ella$_1$.
3:OBJ$_2$ wrote-3sg.PST PREP her.
'Sandra$_1$ called Mónica$_2$ and Rogelio wrote her$_1$.'

Our experiments will investigate whether there is indeed a relationship between reference form markedness and selection of a *non-default* antecedent in Spanish, where markedness is expressed by inclusion of more material, as compared to English where markedness is expressed prosodically. But first, we will describe the two factors, *grammatical role* and *plausibility*, which each seem to play a role in selecting the antecedent of unmarked expressions.

3.1.2 *Grammatical Roles and Antecedent Selection*

There are several viewpoints under which the grammatical role of the potential antecedent has an influence in pronoun resolution. Under the Subject Assignment Strategy (SAS), all pronouns preferentially refer to subjects (Crawley et al., 1990). Pronouns are most often produced as subject, presumably because they most often refer to the topic of the previous sentence, which is itself usually the subject. This is true for both English and Spanish, as the speakers of these languages usually put *Given* information before *New* information (Vallduví and Engdahl, 1996; Arnold et al., 2000). Thus, a pronoun will most often refer to the subject/topic. A number of experiments have found evidence for the subject-assignment strategy in pronoun comprehension (e.g. Crawley et al., 1990).

On the other hand, there are good reasons to expect that speakers and hearers are sensitive to Grammatical Role Parallelism. Parallelism is quite a general phenomenon in language processing. For instance, experiments on syntactic priming in

comprehension have suggested that exposure to a sentence with a given structure facilitates the comprehension of subsequent sentences with a parallel structure(Frazier et al., 1984; Branigan et al., 1995). Along similar lines, the Parallel Function Strategy (PFS) of antecedent selection suggests that Grammatical Role Parallelism facilitates anaphora comprehension (cf. Callahan et al., 2010). Several authors (Grober et al., 1978; Chambers and Smyth, 1998; Smyth, 1994) suggest that, all else being equal, a pronoun should refer to a syntactically parallel antecedent in a foregoing sentence. Under the PFS, a pronoun refers to the entity in a foregoing clause that has the same grammatical role. A subject pronoun will refer to the preceding subject, and an object pronoun will refer to the preceding object.[2]

We have already discussed some empirical evidence in favour of the PFS. Looking at ambiguous subject pronouns, Venditti et al. (2001), for instance, found a preference for unmarked pronouns to be resolved to the subject antecedent. This can be taken as evidence for the SAS, and has been for similar experiments. However, for Spanish, Blackwell (2001, see also Baauw et al. 2011) reports that unmarked clitics preferably refer to the object antecedent. Taken together these suggest that a pronoun is taken to refer to the parallel element. However, these experiments were done in different languages. Our goal is to look at both languages using virtually identical materials and an identical design for comparability.

We have already pointed out that languages may differ substantially in structure. This is true of English and Spanish, since in Spanish, unlike English, different word order rules apply to object pronouns as compared to object full NPs. In the case of object pronouns (e.g. ex. 4), the accompanying clitic is necessarily placed *before* the verb, whereas in an analogous sentence with a full noun phrase, the noun phrase must appear *after* the verb. This alternation creates the interesting situation that Spanish object pronouns can never be strictly linearly parallel with their antecedents, although at a more abstract level

2 The strong version of this hypothesis suggests that grammatical roles need to be completely identical, so that oblique object pronouns would not refer to direct object pronouns. This suggestion was refuted by Kehler et al. (2008), and so we take the weak version, that subjects refer to subjects, and objects refer to objects.

they are both objects. The comparison between English and Spanish thus provides information about how parallel the word order of a structure must be to lead to a preference for a parallel noun phrase.

There is still no consensus on whether Grammatical Role Parallelism plays a role of importance in antecedent selection. Some studies that have supported the SAS, where all pronouns are assumed to preferentially refer to the preceding subject, are not very informative because they only used subject pronouns (where the predictions for SAS and PFS are identical). We used ambiguous object pronouns in both the English as well as in the Spanish experiments, so that we could pit the SAS against the PFS: object pronouns provide a point of critical difference between the PFS and SAS because, according to the SAS, the subject is the preferred antecedent, and according to the PFS, the object is the preferred antecedent. Importantly, the current study also intends to add to this debate by investigating a factor that may have led to equivocal findings in previous studies.

We direct the reader to Kehler et al. (2008), for a discussion of the idea that parallelism occurs as a result of parallel event structure sequential clauses via coherence relations. They suggest that grammatical role effects are not in fact driven by thematic roles or order-of-mention, although this is beyond the scope of the current discussion.

In spite of differences of opinion about how parallelism functions, Kehler et al. (2008) and Smyth (1994) make the same prediction regarding the interaction between parallelism and accent. Smyth specifies this as a rule of the grammar of English, whereas Kehler et al. (2008) suggest the selection of a non-paralel entity occurs because accent indicates non-parallel co-referentiality as a feature of information structure (which is quite compatible with Levinson (2000) and Huang (2000)'s description of referent markedness).

3.1.3 *Plausibility and Antecedent Selection*

The third factor affecting antecedent preference that we will investigate is plausibility. Consider sentences (6) and (7). Structurally oriented theories (e.g. both SAS and PFS) predict no differ-

ence in antecedent preference between (6) and (7), although it is clear that in one case (6) the subject of the first clause should be chosen, and in the other case (7), the object of the first clause should be chosen. A purely structural choice would lead to an extremely implausible, if not impossible, interpretation.

(6) *Alex$_1$ murdered Barry, and Sarah arrested him$_1$.*

(7) *Cam captured Alfred$_2$, and Jane released him$_2$.*

These examples suggest that pronoun reference is ultimately limited by the listener's sense of what is plausible, given world knowledge. However, it is unclear either theoretically, or experimentally, under what conditions a more plausible antecedent becomes the referent, and to what extent the effect of plausibility can be overcome by other factors such as grammatical role or markedness of the reference form.

Although most experimentalists have assumed that plausibility will have some effect, they have paid relatively little attention to it as a contributing factor. Moreover, the earlier work that we cited does not appear to have adequately controlled for the effects of world knowledge. For example, in testing for plausibility, Crawley et al. (1990) asked judges whether both names were plausible antecedents. If the question was indeed so posed, it seems likely to have led the judges to say yes, unless the plausibility differences were very large. Chambers and Smyth (1998) seem not to have controlled for plausibility with any measures beyond the authors' intuition either. This suggests that the differences in outcome may be due to the plausibility of different readings across experiments, rather than the grammatical roles that were being investigated.

With respect to interactions with markedness, Huang (1991) (see also Huang, 2000, Blackwell, 2001, cf. Hirschberg and Ward, 1991) suggest that markedness of form is not as important as, for instance, plausibility in pronoun resolution, and that when there is a plausibility bias, markedness will not play a role in antecedent selection.

3.1.4 *The Present Study*

Our study will look at the effect of reference form (unmarked versus marked) in two plausibility conditions, comparing equiplausible sentences, where both antecedents are plausible, with biased sentences, where only one of two antecedents is plausible. We expect that listeners will interpret marked forms as referring to the alternative, initially less preferred antecedent, regardless of whether the markedness is realized prosodically (English) or morphologically (Spanish). This will especially be the case in equiplausible sentences, where choosing the alternative antecedent leads to a plausible interpretation. However, it is unclear to what extent markedness can persuade listeners to choose an antecedent associated with a less plausible interpretation of the sentence. For ease of exposition, we present the equiplausible conditions of the English study and the Spanish study together as Experiment 1, and the biased English and Spanish conditions as Experiment 2.

In Experiment 1, we looked into the effects of reference form markedness in *equiplausible* sentences, where the choice of either of the two possible antecedents leads to a plausible interpretation. The results of this experiment tell us: 1) whether there is evidence for either SAS or PFS; 2) whether this preference is the same in Spanish, which does not (and cannot) exhibit strict syntactic parallelism, as in English, where pronoun and full noun phrases have the same structure; 3) whether the choice of antecedent that is made in the unmarked condition will be completely reversed by marking the pronoun; and 4) whether this *switch* in Spanish depends on the form of emphasis: pronoun doubling, pronoun doubling with an accented pronoun, or both.

In Experiment 2, we used semantically *biased* sentences, in which one of the antecedents is preferred, because the other is implausible. It is important to note that the alternative antecedent is generally implausible, but not completely impossible with regard to world knowledge. The results of this experiment allow us to comment on whether markedness works to switch antecedent in this kind of sentence too, or whether the implausibility of the alternative (less active) antecedent blocks the switch, which may be the case if plausibility is a more important criterion for resolution than reference form.

3.2 EXPERIMENT 1: EQUIPLAUSIBLE SENTENCES

In Experiment 1a, we examined the effects of reference form markedness on antecedent selection in equiplausible English sentences, where the pronoun could plausibly refer to either the subject or the object of the previous sentence. In Experiment 1b, the parallel investigation was carried out in Spanish. By looking at object pronouns, the unmarked pronoun condition enables us to directly test whether the SAS or the PFS was more predictive of our results. The design also allows us to see directly whether a more marked reference form switches antecedent relative to an unmarked form.

3.3 EXPERIMENT 1A—EQUIPLAUSIBLE SENTENCES: ENGLISH

3.3.1 *Methods*

3.3.1.1 *Participants*

Twenty-seven native speakers of English (nine women) residing in Edmonton, Canada participated in the experiment.

3.3.1.2 *Materials and design*

We constructed a perception experiment with Accent as a within items factor (Appendix A). There were three conditions, presented in examples (8–10):

(8) *Sandra called Monica and Roger emailed her.*
(Pronoun without accent, Pr)

(9) *Sandra called Monica and Roger emailed* HER.
(Accented Pronoun, AccPr)

(10) *Sandra called Monica and* ROGER *emailed her.*
(Accented Subject, AccSubj)

In the Accented Subject condition, the subject (first proper noun) of the second clause was accented. This was included as a distractor condition so that participants' attention was not drawn to the possibility that accent was linked to antecedent switch in pronouns. There were three lists; only one version of each sentence appeared in each list. Participants each heard one list.

The number of sentences with a given accent type was the same between lists.

Proper names used in the first clause had the same gender, so that the pronoun was (morphologically) ambiguous. The proper names used were chosen to be unambiguously male or female. All items were recorded by a native speaker of English. There were 29 sentences in total, presented in a pseudo-randomized manner. As an initial step to making sure that both antecedents were plausible, we used *resemblance-related* sentences (where the two clauses of the sentence described related actions or events, see Kehler, 2002), which were then subjected to two materials tests.

MATERIALS TESTS. We conducted two pretests. In the first task, we measured *pronoun selection bias* with a forced-choice antecedent selection task. In this task, participants identified the *preferred antecedent* of the pronoun of a given sentence by reading each sentence and filling in the name of the antecedent. We term the potential antecedent that was not chosen in the materials test the *alternative antecedent*. The item characteristics are displayed in Table 1. These data already suggest that when both antecedents were relatively equiplausible, the object pronoun was preferentially chosen.

In the second task, we collected *plausibility* ratings for all items (Table 1). The instructions stressed that participants evaluate the stimuli on the plausibility of the two subsequent *events*, rather than the acceptability or plausibility of the two sentences. To this end, each of the two sentences was presented in square brackets. Participants saw one of two versions of each item. In the first version, the pronoun was replaced by the subject antecedent, as in: [*Sarah invited Vivian over for dinner.*] [*Evan gave Sarah the address*]; in the second version, the pronoun was replaced by the object antecedent, as in: [*Sarah invited Vivian over for dinner.*] [*Evan gave Vivian the address.*] Participants rated how plausible each set of events was on a scale from one to seven. These materials tests ascertained that both antecedents were roughly equiplausible.

PHONETIC ANALYSIS. A survey of naturally occurring accented pronouns suggested that L+H*, an accent where the pitch

Table 3.1: Item characteristics, as derived from materials tests.

	EXPERIMENT 1A		
	PLAUSIBILITY		ANTECEDENT PREFERENCE
	SUBJECT	OBJECT	
Sandra called Monica,	4.90	5.50	1.90
And Roger emailed her.	(0.10)	(0.10)	(0.10)

	EXPERIMENT 1B			
	PLAUSIBILITY		ANTECEDENT PREFERENCE	
	SUBJECT	OBJECT	WITH PRON.	NO PRON.
Sandra llamó a Mónica,	4.60	5.20	1.40	1.80
y Rogelio le escribió	(0.20)	(0.20)	(0.00)	(0.00)

	EXPERIMENT 2A		
	PLAUSIBILITY		ANTECEDENT PREFERENCE
	SUBJECT	OBJECT	
SUBJECT ANTECEDENT			
Michelle trained Beth,	6.40	3.00	1.10
and Anne paid her.	(0.10)	(0.30)	(0.00)
OBJECT ANTECEDENT			
Robert registered Joe,	3.10	5.70	1.90
and John taught him.	(0.30)	(0.10)	(0.00)

	EXPERIMENT 2B			
	PLAUSIBILITY		ANTECEDENT PREFERENCE	
	SUBJECT	OBJECT	WITH PRON.	NO PRON.
SUBJECT ANTECEDENT				
Michelle entrenó a Bertha,	6.30	3.70	1.10	1.10
y Ana María le pagó	(0.20)	(0.30)	(0.00)	(0.00)
OBJECT ANTECEDENT				
Roberto registró a José,	4.10	6.00	1.60	1.90
y Gustavo le enseñó.	(0.30)	(0.20)	(0.10)	(0.00)

Note: Mean (Standard Error) Pretest 2 data is labelled as plausibility, and pretest 1 data antecedent preference. The plausibility was rated on a seven-point scale; 1=implausible, 7=plausible. The central value is 4. For Antecedent preference, 1= subject of the first sentence, and 2 = object. A score of 1.90 therefore suggests that the object was chosen 90 percent of the time. The Spanish sentence is a translation equivalent of the English sentence. Pron. = Pronoun.

starts low and rises relatively sharply was the most suited accent for our study (a production study by Hirschberg and Ward, 1991 supports these conclusions).[3] We analyzed all stimuli acoustically ensuring that accented features received an L+H* accent, and that unaccented features did not. Additionally, we measured the *stressed-syllable quartiles* following (Watson et al., 2008). This procedure calls for the pitch to be calculated at the beginning of the pronoun, one quarter of the way through the pronoun, midway through the pronoun, three quarters of the way through, and at the end. This analysis indicated that, compared with the unaccented pronoun, the Fo value of the accented pronoun was higher at the first measurement for the pronoun, one quarter of the way through (first quartile), half way through (median), three quarters, and at the end of the utterance. The analysis, visible in Table 2, also demonstrates that the pitch accented pronouns showed the intonation curve that would be typical of L+H*. In the accented pronoun condition, the pitch dips below the other conditions immediately preceding the pronoun, before eventually rising well above the other conditions.

We performed two-tailed t-tests, comparing the pitch of Pr and AccPr at the five points created as stressed-syllable quartiles. We found that there were significant differences between the two conditions at Q1, Median, and Q3 (p-values < .05). Additionally, there was a difference in the length of the pronoun, with accented pronouns significantly longer than unaccented pronouns (AccPr = .31 ms; Pr = .16 ms; $t(27) = 11.480$, $p < .0005$).

PROCEDURE. Participants listened to the sentences and then indicated the antecedent using a seven-point, Likert-type scale, where the NP1/Subject antecedent was on the leftmost side of the scale, and the NP2/Object antecedent on the rightmost side of the scale. Strong preference for the first mentioned antecedent was indicated by circling the leftmost three, strong preference

3 We examined how readers produced Chapter 1 of *Pride and Prejudice* (which is a text particularly rich in accented-pronouns), where Mrs. Bennet exclaims: "I desire you will do no such thing. Lizzy is not a bit better than the others; and I am sure she is not half so handsome as Jane, nor half so good-humoured as Lydia. But you are always giving HER the preference." (Austen, 94nd). Two analysts examined the prosody of 4 speakers producing the utterance. Three of the four speakers produced an L+H* contour. Accordingly, L+H* was the pitch accent used in the studies described below.

Table 3.2: Prosodic characteristics and naturalness ratings of Experiment 1 items.

	EXPERIMENT 1A					
	FO OF THE PRONOUN					
	BEG.	Q1	MEDIAN	Q3	END	NATURALNESS
Pronoun	99.74	98.16	95.86	95.55	98.14	4.00
(Pr)	(4.28)	(1.68)	(1.43)	(1.30)	(2.85)	(0.20)
Accented Pronoun	99.61	124.80	128.12	102.40	105.41	3.90
(AccPr)	(3.09)	(2.49)	(2.67)	(2.80)	(5.68)	(0.20)
Accented Subject	99.61	93.82	92.33	92.50	94.22	3.20
(AccSubj)	(1.89)	(0.96)	(0.82)	(0.75)	(1.59)	(0.20)
	EXPERIMENT 1B					
	FO OF THE PRONOUN					
	BEG.	Q1	MEDIAN	Q3	END	NATURALNESS
Clitic	–	–	–	–	–	4.10
(Cl)						-0.20
Clitic and Pronoun	93.45	93.15	90.93	92.26	94.03	3.60
(Cl+Pr)	(1.35)	(1.00)	(1.45)	(0.94)	(1.15)	(0.20)
Clitic and Accented Pronoun	104.86	114.96	118.86	113.38	97.58	3.30
(Cl+AccPr)	(1.17)	(2.03)	(2.19)	(2.19)	(2.49)	(0.20)

Mean (Standard Error). Fo from beginning to end of the pronoun for Equiplausible items. The mean Fo was calculated at five points in the pronoun's production in each condition: The beginning, one quarter of the way through production (Q1), in the middle (median), etc. for each accenting condition. Naturalness was judged on a five-point scale with 5 natural and 1 unnatural. There were no measurements for the pronoun in the clitic condition since it was not present. The clitic was not measured phonetically since in all cases it was unstressed and it was considered ungrammatical to stress it.

for the second mentioned antecedent was indicated by circling the rightmost three, and uncertainty was indicated by circling zero (Figure 3.1).

												Intonation				
	Question											Unnatural				Natural
1	Who did Brianna ridicule?	Dallas	3	2	1	0	1	2	3	Nathaniel		1	2	3	4	5

Figure 3.1: Example of the ratings task. Participants indicated referent choice and intonation naturalness by circling their preferences.

Prosodic naturalness ratings from one (*unnatural*), to five (*natural*) were also collected. Participants were asked to go with their first intuition. The naturalness scale was shorter and presented after the antecedent rating to encourage participants to pay more attention to the antecedent rating, which is of greater experimental importance. The instructions were read by participants and explained orally by the experimenter.

3.3.2 *Results*

The data on antecedent selection were recoded, so that choices for the first mentioned antecedent were negative (e.g. -3, -2, -1), and choices for the second-mentioned antecedent were positive (e.g. +1, +2, and +3). We performed Repeated Measures ANOVAs both by participant (averaged over items) and by item (averaged over participants). Accent was a within-participants and a within-items factor. To compensate for violations of the sphericity assumption required for a Repeated Measures ANOVA, we applied the Greenhouse–Geisser correction where appropriate. The original degrees of freedom are reported. In order to reduce variance, Item group and List were included as between-subjects factors in the analysis. Below, η^2 should be understood as *partial* η^2.

ANALYSIS OF PROSODIC NATURALNESS. We analyzed the naturalness ratings in order to ensure that all stimuli were rated as acceptable. The analysis found a significant main effect of Ac-

cent (see Table 2) ($F_1(2, 44) = 18.9$, $p < .0005$, $\eta^2 = .46$, $F_2(2, 50) = 19.5$, $p < .0005$, $\eta^2 = .44$) (Difference in degrees of freedom between the analysis of prosodic naturalness and antecedent choice are due to missing cells for some participants.) Bonferroni post-hoc tests indicated that this main effect was due to a difference between the Accented Pronoun condition (AccPr) and the other two conditions (p-values < .0005); the Accented Subject Noun (AccSubj) and the No Accent condition (Pr) were not significantly different (p-values > .7). All of the conditions were rated above the midpoint of the (five-point) scale, indicating that they were considered quite natural.

ANTECEDENT CHOICE. The outcome of the Repeated Measures ANOVA indicated that there was a significant effect of Accent ($F_1(2, 48) = 78.97$, $p < .0005$, $\eta^2 = .77$; $F_2(2, 50) = 92.45$, $p < .0005$, $\eta^2 = .79$; data shown in Figure 3.2 below). Follow-up analyses indicated that there was a significant difference between the accented pronoun condition on the one hand, and the unaccented pronoun conditions on the other hand (p-values < .0005), but not between Accented Subject and Pronoun without Accent conditions ($p_1 = .62$, $p_2 = .30$). All of the conditions had values significantly different from the midpoint of the scale, zero, but the two conditions where the pronoun was unaccented, Pr and AccSubj were significantly above zero (all p-values < .0005), whereas AccPr was significantly below zero. This indicated that accent switched the antecedent of the pronoun.

3.3.3 *Discussion*

Unmarked object pronouns were found to refer to the second-mentioned character: the object of the previous sentence. This result is consistent with the predictions from the Parallel Function Strategy, and does not support the Subject Assignment Strategy. Antecedent preference switched when a marked form, an accented pronoun, was used: participants preferred the first-mentioned character (the subject). In Experiment 1b, we will see whether the same pattern of results obtains in Spanish, a language where object pronouns are obligatorily pre-verbal, rather

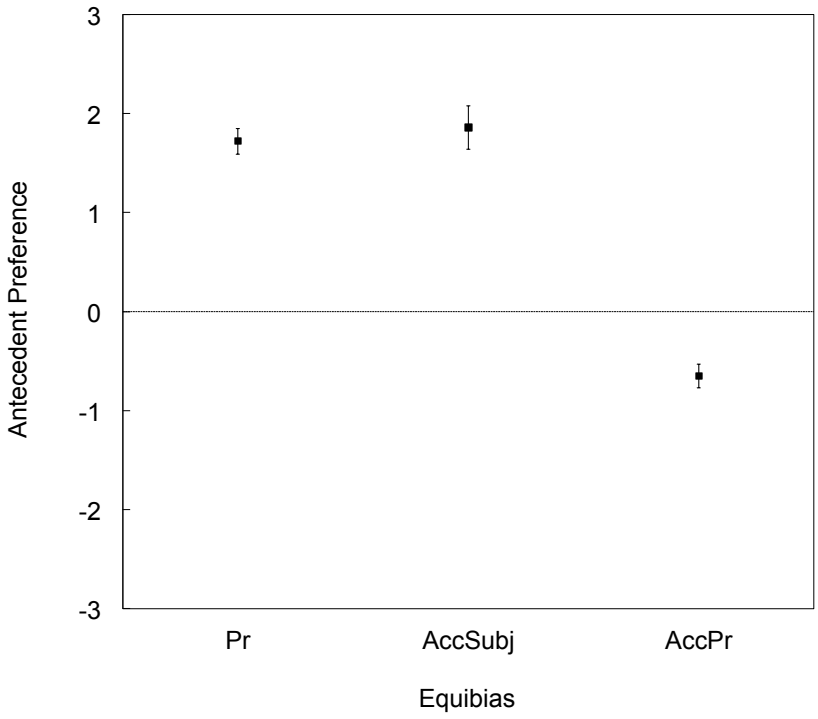

Figure 3.2: Referent Choice for Experiment 1a. Preference for the object is indicated with positive numbers, and preference for the subject with negative numbers. Pr = pronoun. AccPr = accented pronoun. AccSubj = accented subject.

than having a structure that is concretely parallel to the clause containing the antecedents.

3.4 EXPERIMENT 1B—EQUIPLAUSIBLE SENTENCES: SPANISH

Experiment 1a demonstrated that there were strong effects of parallelism in English pronoun comprehension when antecedent plausibility was controlled for. As we discussed in the introduction, Spanish is a particularly interesting test case for the Parallel Function Strategy. Whereas English uses Subject Verb Object (SVO) ordering in both full noun phrase and object pronominal clauses, Spanish does not. The equivalent of an English unaccented pronoun is a preverbal clitic, making the default word

order for Spanish Subject Object Verb (SOV) in cases where there is an object anaphor.

Spanish and English differ not only in terms of where the anaphor is encoded linearly, but also in terms of how anaphors can be made more marked. Whereas English offers only the option of accenting the pronoun, there are two possible Spanish equivalents, the clitic-and-pronoun construction, and the clitic-and-accented-pronoun construction. This experiment tests if the clitic-and-pronoun construction alone produces as large an antecedent switch as in English, and whether accenting the pronoun has an additional effect on antecedent choice.

3.4.1 *Methods*

3.4.1.1 *Participants*

Thirty-eight native speakers of Mexican Spanish (24 women) from the Universidad Iberoamericana community and residing in Mexico City participated in the experiment.

3.4.1.2 *Materials and design*

We created stimuli (Appendix A) with the factor Markedness, which had three conditions (examples 11-13): Clitic–Only (Cl) where there was only an object clitic, Clitic-and-Pronoun (Cl+Pr) for sentences which included the clitic and an object pronoun with no pitch accent, and finally Clitic–and–Accented–Pronoun (Cl+AccPr). [4]

(11) *Sandra llamó a Mónica, y Rogelio le escribió.* (Cl)

(12) *Sandra llamó a Mónica, y Rogelio le escribió a ella.* (Cl+Pr)

(13) *Sandra llamó a Mónica, y Rogelio le escribió a* ELLA (Cl+AccPr)

TRANSLATION PROCEDURE. The English materials from Experiment 1 (including instructions and materials tests) were translated in the following way: first, a certified English–Spanish

4 The translators agreed that subject accenting sounded highly unnatural in Spanish, in contrast with English, where it was considered fairly natural. Inclusion of an unnatural form could fundamentally change the nature of the task, making the two experiments less comparable, thus the accented subject condition was dropped.

translator of Mexican origin translated the materials. In cases where the form of the translation was not possible with an object pronoun, a new item was made. These materials were then examined by a second, English–Spanish translator of Mexican origin. Finally, a native speaker of Mexico City Spanish read the sentences for grammaticality and naturalness once again before recording them.

MATERIALS TESTS. The materials tests were analogous to those performed in English. In the antecedent choice materials test, we included two conditions, one with the pronoun and one without a pronoun (Table 2). The initial results suggest that in written sentences without a pronoun, the object of the preceding sentence was chosen as the antecedent of the pronoun, whereas in sentences with a pronoun, the antecedent is switched. The plausibility rating task that we used for English required a few adaptations necessary for use in Spanish: The stimuli for the plausibility scores took the form *Sandra llamó a Mónica y Rogelio escribió a Mónica (or Sandra)*. The clitic was not included because it is ungrammatical to include a clitic with the noun phrase it is meant to represent. Three items were removed since experimenter/translator intuitions were not found to match with those of the participants in the pretest. These values are not included in the item characteristics. [5]

PHONETIC ANALYSIS OF PITCH. The materials were analyzed acoustically using the same procedure outlined in Experiment 1a. The two pronoun conditions were different in terms of Fo (Table 2). Pitch lowering in the Cl+Pr at the median position is attributable to the word *ella* which was realized with a palatal approximant, leading to an attenuation of the pitch. The intonation of the pronouns is consistent with Spanish Tobi's L+H* in the Cl+AccPr condition, and L* in the Cl+Pr condition (De-la-Mota et al., 2010). Our native speaker's intuitions matched the descriptions of De-la-Mota et al.. for Mexican Spanish: L* is consistent with broad focus in a declarative utterance, and L+H* is consistent with broad or narrow focus in a declarative utter-

5 The word *probabilidad* was used; no direct translation of plausibility exists. *Probabilidad* is closer in meaning to probability, but the translators felt it was the closest translation equivalent in Spanish.

ance (due to a technical error, two items were missing from the Cl+Pr). In total, there were 27 items.

Statistical analysis of the acoustic features showed that there was a significant difference in pitch at the Min, Q1, Median, and Q3 stressed syllable quartiles ($p < .0005$). The accented pronouns were slightly but significantly different in length (Cl+Pr mean = .14 ms, SE = .01; Cl+AccPr mean = .28, SE = .01, $t(25) = 3.50$; $p < .0005$).

PROCEDURE. The procedure was identical to Experiment 1a.

3.4.2 *Results*

Data coding was the same as in Experiment 1a. We analyzed the mean ratings with Repeated Measures ANOVA. Markedness was a within-subjects and within-items factor. The Greenhouse-Geisser correction was applied when the sphericity assumption of the ANOVA was violated; the original degrees of freedom are reported. In order to reduce variance, Item Group and List were included as between-subjects factors in the analysis. Below, η^2 should be understood as *partial* η^2.

ANALYSIS OF PROSODIC NATURALNESS. We found a significant main effect of Markedness ($F_1(2,60) = 15.67$, $p<.0005$, $\eta^2 = .34$; $F_2(2,48) = 50.91$, $p<.0005$. $\eta^2 = .68$). Bonferroni post-hoc analyses indicated that the Cl condition was judged to be more natural than the Cl+Pr and Cl+AccPr conditions (p-values < .0005,), and that Cl+Pr was judged to be more natural than Cl+AccPr in the by items analysis ($p < .0005$). All means were above three, the midpoint of the (five-point) naturalness scale (Table 2).

ANTECEDENT CHOICE. There was a highly significant main effect of Markedness ($F_1(2, 62)=55.42$, $p_1 < .0005$ $\eta^2 = .641$, $F_2(2, 48) = 65.0$, $p_2 < .0005$, $\eta^2 = .73$; shown in Figure 3.3). Bonferroni post-hoc tests indicated that Cl was significantly different from Cl+Pr and Cl+AccPr (all p's <.0005), and Cl+Pr and Cl+AccPr were not significantly different ($p_1 = .11$, $p_2 = .19$). The arithmetic values of both the Cl+Pr and Cl+AccPr conditions were lower

than zero. We tested these values against the central value of the scale, zero, with a one-sample t-test, to test the hypothesis that they were lower than zero. T-tests showed that Cl+AccPr was significantly lower than zero ($t_1(33)$ =3.83, p<.001; $t_2(26)$ = -3.86, p < .001), and in Cl+Pr, the difference was significant in the by-subject condition, and marginally significant in the by-items condition ($t_1(33)$ = -2.48, p < .05; $t_2(26)$ = -1.66, p = .054).

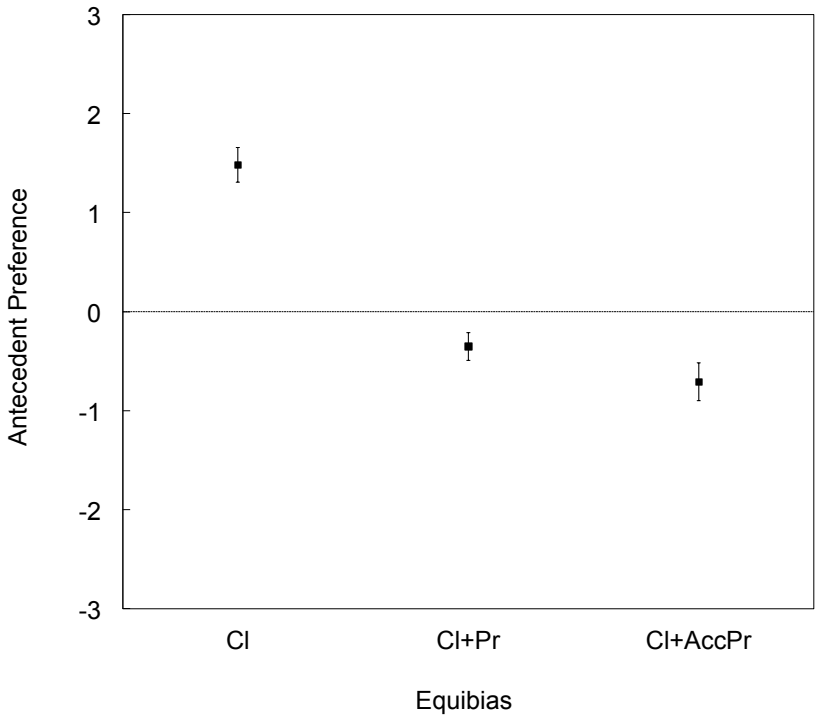

Figure 3.3: Referent choice for Experiment 1b. Preference for the object is indicated with positive numbers, and preference for the subject with negative numbers. Cl = clitic. Cl+Pr = clitic and pronoun. Cl+AccPr = clitic and accented pronoun.

3.4.3 *Discussion of Experiment 1*

Experiment 1 tested the effect of reference form on antecedent selection in equiplausible sentences. First of all, in both English and Spanish, unmarked object pronouns were preferentially linked to the grammatical object of the preceding clause. This is according to the predictions of the PFS. It was uncertain whether Spanish would show equally strong effects of the PFS, since in Spanish the linear ordering of object anaphors is quite different from that of the clause containing the potential antecedents; the results suggest that the more abstract functional parallelism is the important element. These results are inconsistent with the SAS hypothesis, as the subject of the preceding clause is only chosen in a small minority of cases.

The use of a more marked reference form resulted in the choice of the alternative, initially unpreferred antecedent in both languages. In English, the accented object pronoun is interpreted as referring to the subject of the previous sentence. In Spanish, we see essentially the same effect: in both marked conditions, the *clitic plus unaccented pronoun* and *the clitic plus accented pronoun*, the subject is chosen as the most probable antecedent. There was no significant difference between the two marked conditions in Spanish, suggesting that adding emphasis by accenting the pronoun does nothing to add to the "force" of the markedness.

In summary, this experiment suggests that in the absence of a plausibility bias, reference is determined by the interaction of reference form and parallelism in terms of grammatical role. In Experiment 2, we examine how this interaction takes shape when there is a strong bias toward one of the potential antecedents.

3.5 EXPERIMENT 2

The results of Experiment 1 suggest that when antecedents are equiplausible, object pronouns preferentially refer to the object of a previous sentence. This is true for both English and Spanish, irrespective of the differences between these languages in terms of object pronoun–verb ordering. The marked-

ness manipulation revealed that more marked forms led participants to choose subject antecedent, in contrast to unmarked forms, which led to the choice of an object antecedent. What remains to be seen is to what extent plausibility can override these effects. Under some accounts of pronoun resolution, language users will always eventually choose the most plausible antecedent, regardless of markedness or other factors (e.g. Huang, 2000). On the other hand, it seems that language must also offer the opportunity to describe an event in the real world that is not plausible at all (e.g. *John witnessed Bill rob a bank, but then* JOHN *was arrested by mistake*). Under that view (cf. Kehler, 2002), less plausible antecedents can be preferred, but their unlikeliness must be indicated by linguistic means such as emphasis.

3.6 EXPERIMENT 2A—BIASED SENTENCES: ENGLISH

3.6.1 *Methods*

3.6.1.1 *Participants*

The same 27 participants from Experiment 1A took part in this experiment.

3.6.1.2 *Materials and design*

In this experiment, we tested the effect of accent on pronouns with a biased interpretation – where world knowledge creates a preference for one of the available antecedents over the other. In order to create this bias we used *result* sentences according to the classification proposed by Kehler (2002). Result-related sentences often appear in the literature as subordinate *because* clauses, however, Kehler (2002, inter alia) specifies that the important distinction between Result and Resemblance is not a syntactic difference, or even that there are external markers indicating the relations between the clauses. Instead, it is a pragmatic distinction; if the comprehender perceives the action in the first utterance as the cause of the events in the second utterance, they are result-related. Connectives and syntactic struc-

ture may influence the relation, but they are not necessary to determining the coherence relation.

In our materials, there was either a bias to the subject of the first clause (Subject-biased pronouns), or to the object of the first clause (Object-biased pronouns; for materials, see Appendix B). As in Experiment 1a, there were three accent conditions: No Accent (Pr), Accented Subject (AccSubj) and Accented Pronoun (AccPr) presented as a within-participants and within-items factor.

MATERIALS TESTS. The materials tests from Experiment 1a were carried out concurrently with stimuli of Experiment 2a (reported in Table 1). On the basis of an antecedent selection pretest, 16 items were selected for the Subject-bias condition, and 12 items for the Object-bias condition (some of the items changed category after the materials test).

PHONETIC ANALYSIS. Acoustic analysis was carried out on these items in the same manner as in Experiment 1a. The materials used in this experiment had the same pitch accents as in Experiment 1a, L+H* and L* (Table 3). Two-tailed t-tests comparing the mean stressed syllable quartiles found that there was a significant effect of accent at Q1, Median, and Q3 (p-values < .05.). There was also a significant difference in the length of the pronoun, with accented pronouns longer than unaccented pronouns (AccPr mean = .28, SE = .01; Pr mean = .14, SE = .01; $t(26) = 10.09$ $p < .0005$).

3.6.2 *Results*

ANALYSIS OF PROSODIC NATURALNESS. As in Experiment 1, we performed Repeated Measures ANOVAs on the naturalness ratings, with Bias (*Subject-bias* versus *Object-bias*) as a within-participants and between items factor, and Accent (*no accent*, versus *subject accent*, versus *pronoun accent*) as a within-participants and within-items factor. There was no main effect of Bias, but there was a significant main effect of Accent ($F_1(2, 44) = 31.68$, $p < .0005$, $\eta^2 = .59$, $F_2(2, 44) = 60.31$, $p < .0005$, $\eta^2 = .73$). This main effect of Accent was qualified by a significant

Table 3.3: Prosodic characteristics and naturalness ratings of Experiment 2 items.

	EXPERIMENT 2A						
	FO OF THE PRONOUN					NATURALNESS	
	BEG.	Q1	MEDIAN	Q3	END	SUB. BIAS	OBJ. BIAS
Pronoun	102.91	104.27	101.28	98.40	99.29	4.20	4.40
(Pr)	(4.30)	(1.72)	(1.28)	(1.17)	(2.03)	(0.10)	(0.10)
Accented Pronoun	96.57	116.25	135.62	119.85	109.27	3.30	3.00
(AccPr)	(1.67)	(6.30)	(2.59)	(3.62)	(4.84)	(0.20)	(0.3)
Accented Subject	96.51	102.06	98.20	96.57	95.16	3.90	4.20
(AccSubj)	(3.45)	(1.19)	(1.06)	(0.87)	(1.27)	(0.20)	(0.20)
	EXPERIMENT 2B						
	FO OF THE PRONOUN					NATURALNESS	
	BEG.	Q1	MEDIAN	Q3	END	SUB. BIAS	OBJ. BIAS
Clitic	–	–	–	–	–	4.40	4.30
(Cl)						(0.20)	(0.20)
Clitic and Pronoun	94.63	94.37	92.19	93.57	94.97	3.90	3.90
(Cl+Pr)	(0.90)	(1.02)	(1.38)	(1.62)	(2.33)	(0.10)	(0.20)
Clitic and Accented Pronoun	101.23	109.59	114.12	111.89	98.80	3.60	3.40
(Cl+AccPr)	(0.91)	(1.67)	(2.04)	(3.73)	(7.53)	(0.20)	(0.20)

Mean (Standard Error). Fo from beginning to end of the pronoun for Equiplausible items. The mean Fo was calculated at five points in the pronoun's production in each condition: The beginning, one quarter of the way through production (Q1), in the middle (median), etc. for each accenting condition. Naturalness was judged on a five-point scale with 5 natural and 1 unnatural.

interaction of Bias and Accent ($F_1(2, 44) = 5.43$, $p < .01$, $\eta^2 = .20$, $F_2(2, 44) = 5.86$, $p < .01$, $\eta^2 = .21$), (Table 3). Follow-up analyses revealed that Subject-bias AccPr were more natural than Object-bias AccPr, whereas conversely, Subject-bias SubjAcc was less natural than Object-bias SubjAcc. None of the conditions were rated unnatural.

ANTECEDENT PREFERENCE. The data on antecedent selection were recoded as in Experiment 1 for purposes of graphical display (data shown in Figure 3.4). Choices for the subject were negated (e.g. -3, -2, -1), and choices for the object remained positive (e.g. +1, +2, and +3). For statistical analysis, we used a different coding, as we wanted to compare strength of bias, regardless of whether this bias was directed towards subjects or objects. This specific recoding was done as follows: for the object-bias condition, choices for the subject became negative (e.g. -3, -2, -1), and choices for the object remained positive (e.g. +1, +2, and +3). For the subject-bias condition, it was the other way around: choices for the first-mentioned antecedent (subject) remained positive (e.g. +1, +2, and +3), and choices for the second-mentioned antecedent (object) became negative (e.g. -3, -2, -1). Positive numbers thus signify a preference for the *normally* preferred antecedent (i.e., subject in subject-biased sentences, and object in object-biased sentences).

The Repeated Measures ANOVAs did not produce a main effect of Bias (F-values < 1.5). They did, however, show a main effect of Accent ($F_1(2, 48) = 31.81$, $p < .0005$, $\eta^2 = .57$; $F_2(2, 44) = 25.20$, $p < .0005$, $\eta^2 = .53$). Post-hoc analyses indicate that this effect was driven by differences between AccPr and the two other conditions (p-values < .0005). This main effect is qualified by an interaction between Bias and Accent which is significant in the by-items analysis, and marginally significant by participants ($F_1(2, 48) = 2.61$, $p = .10$, $\eta^2 = .09$; $F_2(2, 44) = 4.28$; $p < .05$, $\eta^2 = .16$). The locus of trend towards an interaction appears to be due to a difference in the AccSubj items, since Subject-bias sentences with accented subject NPs tended to show higher antecedent preference than Object-bias sentences with accented subject NPs (significant in the F_2 analysis, but not the F_1 analysis). No other post-hoc results were significant ($p > .26$).

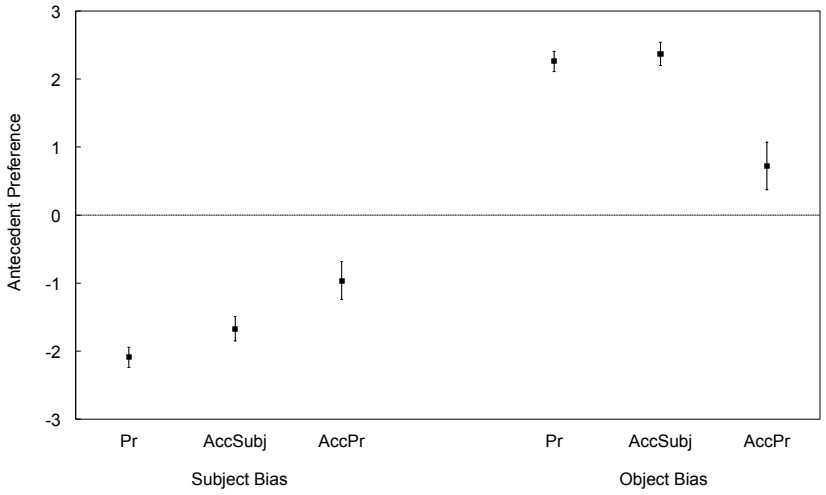

Figure 3.4: Referent choice for Experiment 2a. Preference for the object is indicated with positive numbers, and preference for the subject with negative numbers. Pr = pronoun. AccPr = accented pronoun. AccSubj = accented subject.

Paired t-tests comparing the conditions to the central value of the scale, zero, showed that there was no evidence of antecedent switch in either Object-biased or Subject-biased pronouns: all means were significantly different from zero (p-values < .0005).

3.6.3 *Discussion*

The outcome of Experiment 2a illustrates the strength of the plausibility bias in antecedent choice. In the cases where the object pronouns were unmarked, plausibility determined pronoun antecedent identity. Although having a more marked reference form failed to switch the ultimate choice of antecedent, it did affect the strength of preference for an antecedent. In both Subject-bias and Object-bias conditions, antecedent preference ratings were significantly attenuated when the pronoun was accented; more marked forms pushed values toward the centre of the scale. In other words, accenting the pronoun apparently made participants less certain whether they had made the right choice, although plausibility remained the main criterion. In the next section, we will see whether these results for English sentences hold true for Spanish as well.

3.7 EXPERIMENT 2B—BIASED SENTENCES: SPANISH

Experiment 2a on English provided evidence that plausibility overrides factors such as parallelism or reference form markedness. Experiment 2b tests whether Spanish behaves in the same way.

3.7.1 *Methods*

3.7.1.1 *Participants*

The same participants that took part in Experiment 1b took part in experiment 2b, concurrently.

3.7.1.2 *Materials and design*

MATERIALS TESTS. The materials tests were identical to and performed concurrently with those performed in Experiment 1b (Table 1). There were 16 items in the Subject antecedent condition and 14 in the Object antecedent condition, assigned on the basis of the materials tests (Appendix A).

PHONETIC ANALYSIS. As above, we analyzed each item individually, confirming that Cl+AccPr stimuli showed an L+H* pattern, and the Cl+Pr condition showed an L*. We performed two-tailed t-tests comparing Cl+Pr and Cl+AccPr, finding that there were significant differences at Minimum, Q1, Median, and Q3 (p-values$<$ .0005). The overall pattern is shown in Table 3. As above, L+H* accented pronouns were observed to be longer than L* accented pronouns (Cl+Pr mean = .24 ms, SE = .01; Cl+AccPr mean = .30 ms, SE = .01; $t(31) = 3.21$, $p < .005$)

PROCEDURE. The procedure was the same as in Experiment 1b.

3.7.2 *Results*

PROSODIC NATURALNESS. We conducted Repeated Measures ANOVAs taking Bias (Subject-bias versus Object-bias; shown in Table 3) as a between-items and within-participants factor, and

Markedness as a within-items and participants factor (Cl, Cl+Pr, Cl+AccPr). There was no significant main effect of Bias. There was, however a significant main effect of Markedness ($F_1(2,58)$ = 18.56, p< .0005, η^2 = .39; $F_2(2,48)$ = 25.78, p< .0005, η^2 = .52). Post-hoc analyses suggest that this follows from a pattern of Cl being rated more natural than Cl+AccPr ($p_1 < .0005$, $p_2 < .0005$) and Cl+Pr ($p_1 = .0005$, $p_2 < .0005$), and Cl+Pr being rated significantly above Cl+AccPr in the by-items analysis ($p_1 < .092$; $p_2 < .05$).

The main effect of Markedness was qualified by an trend of Markedness and Bias in the by-subjects analysis, ($F_1(2,58)$ = 2.97, p = .06, η^2 = .09; $F_2(2, 48)$ = .773, p =.46). This interaction was due to a tendency for the Cl+AccPr condition to be higher rated in the Subject-bias condition than in the Object-bias condition in the by-object analysis ($t_1(32)$ = 1.56, p =.13, $t_2(28)$ = 1.90, p = .07). There was a trend in Cl for Subject-bias to be rated above Object-bias in the by-subject analysis, but not the by-item analysis ($t_1(32)$ = 1.78, p =.08, $t_2(28)$ = .20, p = .84). Cl+Pr did not differ between Subject- and Object-biased items (Cl+Pr $t_1(31)$ = .86, p = .40 ; $t_2(28)$ =.21 , p = .84, n.b. degrees of freedom in the subject analysis are lower due to data loss).

REFERENT CHOICE. Data were transformed in the same way as Experiment 2a (see Figure 3.5 for untransformed values). Repeated Measures showed a significant effect of Bias in the by-subject analysis, and a trend in the by-items analysis $F_1(1,31)$ =18.42, p<.0005, η^2 = .37; $F_2(1,24)$ =2.609, p_2 = .12, η^2 = .098, and a significant effect of Markedness ($F_1(2,62)$ =13.40, p<.0005, η^2 = .302; $F_2(2,48)$ =5.714, p_2 < .01, η^2 = .192). These main effects were, however, qualified by an interaction between Bias and Markedness, which was significant in the by-participant analysis, and marginally significant by items ($F_1(2,62)$ = 7.538, p<.005, η^2 = .196; $F_2(2,48)$=2.607, p = .08, η^2 = .09). Follow-up analyses showed that Cl+AccPr was significantly lower in Object- than in Subject-bias in the by-participant analysis, and marginally so by items (t_1 (33) = 4.384, p_1 < .0005, t_2 (28) = 1.82 p_2 =.080)—a pattern that was evident as a tendency for Cl+Pr ($t_1(33)$ = 1.969, p_1 = .057, $t_2(28)$ = 1.17, p_2 = .11), but not in Cl ($t_1(33)$= .599, p_1 = .55., $t_2(28)$ = .12, p = .90). In other words, these follow-up tests suggest that the interaction is due to changes

in antecedent choice in the Cl+Pr and the Cl+AccPr conditions that are more pronounced in the Object-bias condition than in the Subject-bias condition.

One-sample t-tests showed that all conditions were significantly different from "0", the central value of the scale (p-values < .0005), except for the object-biased Cl+AccPr condition ($t_1(33) = 1.741$, $p = .091$, $t_2(13) = 1.274$, $p = .113$).

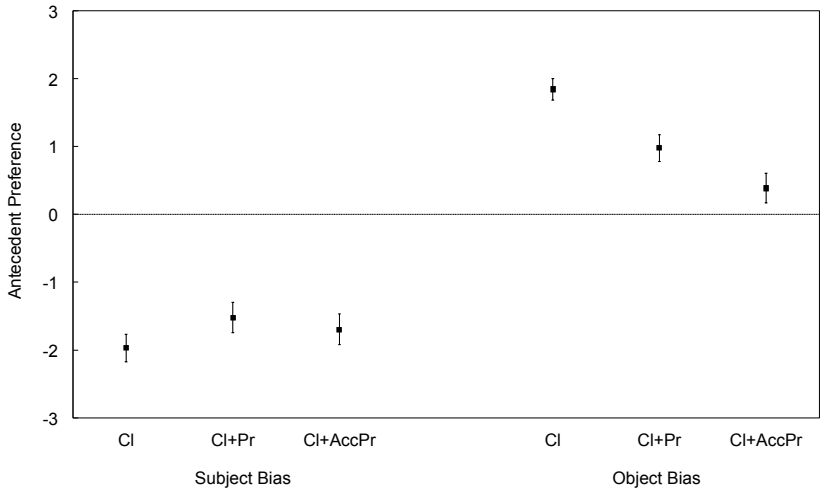

Figure 3.5: Referent choice for Experiment 2b. Positive values indicate the preference for the object, negative values for the subject. Cl = clitic. Cl + Pr = clitic and pronoun. Cl + AccPr = clitic and accented pronoun.

3.7.3 *Discussion of Experiment 2*

Experiments 2a and 2b showed that in both English and Spanish, plausibility is very influential in determining the antecedent of a pronoun. In our stimuli, it was the strongest factor, as we did not find evidence for either an effect of grammatical role or an actual switch of antecedents. However, in both languages, antecedent choice interacted with markedness. In English as well as in Spanish, using a marked reference form significantly attenuated the extent to which participants were certain

of their choice for a specific antecedent. There was one condition, which seemed to produce the largest change in antecedent choice, and that was in the Spanish experiment when the clitic was accompanied by an accented pronoun (Cl+AccPr). This suggests that accenting has an effect on antecedent choice over and above clitic doubling, although this was admittedly only a trend, as the difference with unaccented pronouns (Cl+Pr) failed to reach significance (p-values > .06).

3.8 GENERAL DISCUSSION

In two experiments, we investigated how the form of a referring expression influences antecedent selection. In the first experiment, where we factored plausibility out, we observed an almost exclusive preference for unmarked object pronouns to refer to the grammatical objects of the previous clause. This strong preference was pushed to subject preference when the object pronoun became marked by accentuation (English), or by pronoun doubling with or without pronoun accentuation (Spanish). These findings make two important suggestions.

First of all, these findings support the parallel function strategy (PFS), and are generally inconsistent with the idea that listeners use a subject assignment strategy (SAS) in cases of ambiguity. The strictest interpretation of the PFS holds that, for example, an indirect object would refer to an indirect object in a previous clause, and not, for instance, to a direct object. Two aspects of our data relate to this interpretation. First, in our stimuli, "object" referred to indirect objects, direct objects and oblique objects, which were not matched between clauses. However, being an object *in general* was enough for the PFS to control pronoun resolution, and the type of object appears not to have been of primary importance. The Spanish stimuli especially rule out a strict interpretation of syntactic parallelism, since *noun phrases* referring to humans, as in the first clause of the stimulus sentences, can never be direct objects (e.g. they are obligatorily indirect objects), but *clitics* referring to humans, as in the second clauses of a Spanish stimulus, can be direct objects. In order to create sentences with two potential antecedents, the sentences almost all contained an animate object. Nevertheless,

in Spanish, as in English, the unmarked anaphoric forms were taken to refer to the object antecedent in the first sentence.

These findings thus indicate that the Parallel Function Strategy has to be reformulated, perhaps even in non-syntactical terms such as the order of mention of entities (cf. Järvikivi et al., 2005). For instance, the first mentioned anaphoric expression in a given clause might refer to the first mentioned antecedent in the preceding clause, and, by contrast, the last mentioned anaphor would then refer to the last mentioned (most recent) antecedent in the preceding clause.

Another important finding from the present study is that markedness can cause the listener to select an initially unpreferred antecedent. In both Spanish and English, inclusion of a more marked form resulted in the selection of the alternative antecedent. With respect to how pronouns are marked in Spanish, we found that it did not make a difference whether the doubled pronoun was accented or not: accentuation did not result in a significant increase in preference for the antecedent that was not chosen by the unaccented pronoun. Thus, for antecedents to switch, it is sufficient for a referring form to be marked, regardless of whether this occurs through accentuation (which is the preferred way for English), or through elaboration of form (as in unaccented Spanish pronoun doubling), or both (Accented Spanish pronoun doubling).

In the second experiment, where one potential antecedent was more plausible than the other one, markedness also affected antecedent choice, but not to such an extent that it could switch antecedents as it did in Experiment 1. In all markedness conditions, participants still preferred the more plausible antecedent, although this preference became very small in the accented pronoun condition in object-bias sentences in Spanish; there, a considerable number of participants seemed to also consider the alternative, less plausible referent. Thus, it seems that Huang (2000) was right in claiming that world knowledge information can block the choice for a less plausible antecedent.

But even if markedness was not powerful enough to completely overrule the preference for the most plausible referent, it did significantly change the strength of this preference. So, in this sense, (Kehler, 2002) was also right in suggesting that accentuation is one of the means of saying things that may not be

very plausible. Indeed, whether emphasis was added by pronoun doubling, accentuation, or both, all these options were shown to attenuate the preference for the most plausible referent.

Both experiments thus find strong effects of markedness. These findings suggest that whenever the realization of linguistic elements is more elaborate than is necessary under the specific communicative circumstances at hand, it will act as a signal to the listener to go beyond the default, standard mode of comprehension, and search for a different level of meaning (cf. Levinson, 2000; Huang, 2000). This hypothesis raises some interesting questions, such as: Which tools do different languages have to indicate markedness? Do all means of indicating markedness have the same effect? To what extent do different instantiations of markedness have an additive effect when combined? Also, is it enough to specify that marked forms are less frequent and longer, or is there something special about a high accent and the addition of more material cross-linguistically. This hypothesis appears to make the unintuitive prediction that a low accent, creaky voice, or even a reduced but infrequent form could be used to refer to a less obvious antecedent, a claim that needs further investigation.

It could be concluded from this study that accent is only used to select a normally dispreferred antecedent. However, Chafe (1974) suggests that accent can also be used to warn the listener of something unexpected. In line with this, Dimitrova et al. (2009), found that when people were asked to describe a *blue* banana they saw on a computer screen, the adjective received more accent than when they saw a normal, *yellow* banana, even in experimental conditions where only the noun should be accented (i.e., in a situation where blue lemons and blue bananas were contrasted). This prosodic marking of semantically unexpected information may have a communicative function, in that the speaker appears to be warning the listener that this information is special and needs special attention (see also Wang et al., 2011). This suggests that semantically unexpected antecedents could – or perhaps should – be referred to with marked pronouns.

3.9 CONCLUSION

To summarize, a set of general principles of antecedent comprehension can be defined based on these results. The results suggest that in both English and Spanish, for our experimental manipulations, plausibility outweighed syntax. We interpret this as evidence against the strong forms of syntactic preference heuristics, which would predict that either a first-mention bias or parallelism would outweigh any other biases. When there is a strong bias toward an antecedent, this bias can determine pronoun resolution. Markedness (in the form of prosodic or morphological prominence) was able to lower the strength of the listener's preference in the face of this strong plausibility bias, but could not overturn the preference. When antecedents are equally plausible, however, markedness determined the eventual choice of antecedent. In unmarked cases, some form of the parallel function strategy is at work, such that pronouns appear to take an antecedent from the previous sentence with a comparable grammatical role (in a broad sense). On the basis of our experiments we were able to order factors with respect to their importance for antecedent selection. This does not of course say much about the order in which these factors are used in online pronoun resolution. More research in this area is definitely needed.

4

ACCENTED PRONOUNS AND UNEXPECTED REFERENTS

They didn't expect it of HIM.

ABSTRACT

Past psycholinguistic studies have shown that world knowledge and the markedness of a referring expression are taken into account when comprehending ambiguous pronouns. This study uses ERP to extend past results by examining the interaction of world knowledge-based expectedness and accentuation on the interpretation of *un*ambiguous pronouns. It has been found that speakers sometimes add an accent to intended but implausible or unexpected words, but it is not clear whether listeners can use this information to their advantage. In the current study, we examine how accent contributes to the comprehension of a pronoun referring to an implausible character. We manipulated the plausibility of a character in the context preceding the target sentence so that that character was either expected or unexpected in a following proposition, and the subsequent pronoun referring to that character was either accented or not. Implausible sentences where the pronoun was unaccented produced an early P600 effect, rather than an N400 associated with difficulty in comprehension. Accenting the pronoun produced an Nref effect, reflecting the effort associated with trying to carry out a reference switch (but not succeeding). Alternatively, it is also possible to interpret this effect as a reduced P600, perhaps indicating that interpretation did become easier by having an accent on the pronoun.

4.1 INTRODUCTION

Establishing reference is an integral part of language comprehension, and is essential for listeners to be able to keep track of who is doing what to whom. Once a character is introduced, its representation is active and can be referred to with a pronoun. When a character is less active more information is necessary to refer to that character, so the level of activation of a referent also correlates with prosodic features such as pitch accent: If a character is new in the discourse, or is deemed to be rather unexpected by the speaker, the expression referring to this character is frequently accented. In the present study, we will use ERPs to investigate 1) how comprehenders deal with a pronoun referring to unexpected referents, and 2) whether the processing difficulty elicited by this unexpectedness can be overcome or ameliorated by accenting this pronoun.

We know from previous research that accenting a pronoun can induce a switch of antecedent preference, relative to unaccented words. In a sentence like *John called James, and Sarah emailed him*, comprehenders understand the pronoun as referring to the most active *James*, but when the pronoun is accented, it refers to the less active and thus less expected *John* (Taylor et al., 2013). Interestingly, this effect seems only to be observable with pronouns, and it is difficult to see how accenting a content word could have this same effect. Until now, experimentation on accented pronouns has been restricted to this switching phenomenon to the exclusion of other uses of accenting on a pronoun. But pointing to a less expected, less active antecedent isn't the only reason why a pronoun might be unexpected or why it might be accented.

Recent work in the production literature suggests an alternate, possibly related use of accenting. For instance, speakers can use accenting to draw attention to something unexpected based on world knowledge (Dimitrova et al., 2009). Dimitrova et al. showed that if participants have to describe a pair of fruits with unnatural colours (say a blue banana and a blue orange), they tend to put accent on the colour adjective (e.g. BLUE banana), even though colour is not a contrastive feature (as it

would be with a blue banana and a yellow banana). So speakers may accent something in order to draw attention to information that may be difficult for comprehenders to accommodate in their discourse model. Native speakers of Dutch found this sentence grammatical. The main question of the present experiment is not whether application of an accent to a pronoun with an unexpected referent is grammatical, but rather how does applying an accent to a pronoun with an unexpected referent affect interpretation. Putting an accent on the verb is one means of indicating that the *predicate* or *action* of the verb is odd. Accenting the pronoun, however, should put greater emphasis on the unexpectedness of the *referent* rather than the action.

Example (1) contains an accented pronoun where no reference switch is possible:

(1) A: *Sarah lost the chess match.*
B: *Oh?*
A: *The teacher congratulated* HER.

The use of a pronoun is felicitous since the discourse entity *Sarah* is *activated* since it was introduced into the discourse at an earlier point (Lambrecht, 1996; Arnold and Griffin, 2007; Ward et al., n.d.). But even though *Sarah* has only just been mentioned, and is the only available antecedent, she is an *unexpected* antecedent: if all we know about *Sarah* is that she lost the chess tournament, based on world knowledge we would think it unlikely that the teacher would congratulate her. *Sarah*'s role as a congratulatee is thus unexpected. There is little or no work examining how this *unexpectedness* accent affects language comprehension. Research looking at effects of accent on noun phrases has found that, normally, accenting leads to an increase in the extent of whatever processing would normally take place, for example, if there is normally an N400, the N400 would be even greater with accenting (e.g. Wang et al., 2011). However, in the case of unexpectedness accent, a reference may be pragmatically infelicitous if the pronoun is *not* accentuated.

In the experiment presented below, we manipulated referent expectedness to see whether accenting influences the process-

ing of a pronoun expressing an unexpected referent. Our study adds an important piece to the puzzle by manipulating expectedness while keeping activation level of the discourse entity constant. Previous studies have examined sentences where pronouns were unexpected because no potential characters were available (Osterhout and Mobley, 1995), or where there was a strong initial bias for a character based on syntactic structure, topic structure, or pragmatics (Streb et al., 1999; Van Berkum et al., 2007; Ferretti et al., 2009). However, in these past studies the character violating the bias is not in fact implausible or unexpected, merely *less* expected. In our study, in contrast, the character is highly active, as in (1) where *Sarah* is mentioned as the topic of the first sentence, but nevertheless unexpected, since people losing a tournament are normally not congratulated.

During ERP recording, participants can attend to auditory stimuli naturally, without having to respond to (parts of) these stimuli, and without having to perform any unnatural task. In addition, ERPs allow the intensity of brain function to be sampled directly. Moreover, ERPs provide rich data, including information about latency, scalp distribution and polarity, that can be related to previous work on discourse structure, pragmatics, the acoustic phonetics of accent, and pronoun comprehension.

The aim of the present study can be broken down into two separate issues: What is the effect of an unexpected yet highly salient referent on pronoun comprehension? And what effect does accenting the pronoun with the unexpected referent have on the comprehender?

4.1.1 *Comprehension of an Unexpected Referent*

Before we can consider what effect the unexpectedness accent might have, we need to know how an unexpected pronoun is processed. Semantic and real world knowledge anomalies have been thoroughly examined with ERPs, and, in fact, there have even been a number of studies on comprehension difficulties in pronouns.

Unexpected referents cause difficulty for comprehenders in terms of violations of world knowledge. The most famous ERP

finding associated with this type of anomaly is the N400 effect, for *spread his warm bread with socks* vs. *spread his warm bread with butter* (Kutas and Hillyard, 1984). When comparing the sentences *The teacher congratulated the* loser to *The teacher congratulated the* winner, the word *loser* would show a negative peak 400 ms after the word was presented. Under the classic "integration view", N400 amplitude reflects the ease with which a word is integrated into the existing semantic representation of (part of) a sentence. Semantically anomalous words are harder to integrate, and thus elicit an increased N400 (Brown and Hagoort, 1993). Recent reviews have questioned whether the N400 is related to integration (Kutas and Hillyard, 1984; Brouwer et al., 2012), proposing that N400 amplitude instead reflects the ease of retrieval of lexical information from semantic memory. Since the retrieval of the concept of *loser* is less primed or facilitated than *winner* by *congratulate*, an N400 effect is expected to appear. An interesting question is what happens with pronouns that have unexpected referents, since the concept that the pronoun refers to is not primed or facilitated by the immediately preceding context, although these concepts are "activated" by the preceding global discourse. Since one source of information predicts that the referent is active while another may render it less active, it is an open question whether or not pronouns with unexpected referents will give rise to an N400 effect under the retrieval account. However, an N400 effect is predicted if N400 amplitude is sensitive to ease or difficulty of integration.

At present, we are unaware of any studies reporting N400 effects on *anomalous* pronouns. Most studies in which pronouns violate linguistic rules have generated a P600 instead. For instance, Osterhout and Mobley (1995) showed participants sentences like *The aunt heard that she*, where the *aunt* is a possible referent, and *The aunt heard that he*, where there was no active male referent whatsoever. Osterhout and Mobley found that participants who judged the sentences with mismatching pronouns to be ungrammatical showed a positivity relative to the gender-matched sentences, which was later recognized as a P600 (see also Nieuwland and Van Berkum, 2006).

The P600 has also been shown in a number of cases where, for various reasons, the most accessible referent is not the one referred to. While the P600 in Osterhout and Mobley (1995) in-

dexed a violation, further studies have found P600s for simply going against a specific preference. For instance, Van Berkum et al. (2007) found a P600 effect when comparing a sentence where the gender matched the bias, to a sentence where the gender did not match the bias (*Lisa inspired David because he...*, where normally the bias is toward *Lisa*). This suggests that some combination of knowledge of the world and previous experience with language causes the listener to anticipate reference to a particular character. In another study, Ferretti et al. (2009) found (among other things) a P600 effect in sentences where the antecedent was inconsistent with the preference based on the event-structure of the first clause (e.g. *Sue handed a timecard to Fred. She/?He asked about the upcoming meeting.*).

Under some views, the P600 amplitude reflects effortful syntactic processing (e.g. Hagoort et al., 1993; Osterhout and Holcomb, 1995; Kaan and Swaab, 2003; Hoeks et al., 2004). However, several authors have recently proposed that P600 effects can be elicited by semantic violations, and need not be syntactic in nature (Burkhardt, 2007; Kuperberg, 2007; Brouwer et al., 2012). Indeed the P600s associated with pronouns seem quite prototypical of this "semantic" P600 explanation. From this viewpoint, the unexpected pronouns in this experiment may elicit a P600. That being said, it is important to note that these studies did not actually include a pronoun with a clear real world anomaly. Besides a P600 effect, Ferretti et al. (2009) also identified an *Nref* effect for unexpected pronouns, an effect that has also frequently been reported for difficulty in resolving pronouns. This effect was first reported by Van Berkum et al. (1999). It was elicited by a referring expression for which there was more than one possible antecedent. Thus, the Nref was interpreted as reflecting the search for a correct referent in a case of ambiguity. However, it is not clear whether ambiguity is really necessary for finding an Nref effect. Streb et al. (1999) investigated structural bias by comparing syntactically parallel pronouns to non-parallel pronouns, and found that non-parallel pronouns created what Streb, et al. refer to as an N400-like negativity in the 270 to 420 ms time window, accompanied by another negativity in the 510 to 600 ms time window, both with a largely frontal distribution. Retrospectively, both the temporal and topographical distributions seem consistent with an Nref,

although as in Feretti et al., there was no actual ambiguity. Our unexpected condition could thus also elicit an Nref effect.

In summary, then, the literature suggests that there are three possible ERP responses when comparing pronouns referring to unexpected referents with pronouns referring to more expected words. First, pronouns with unexpected referents could give rise to an N400 effect, especially under the integration view of the N400. However, no empirical study on pronoun processing to date has reported N400 effects. It is therefore more likely that such a comparison will produce a P600 effect. This fits well with the hypothesis that P600s result from difficulty in integrating a character into the discourse model. Under this view the only distinction between pronoun manipulations and other P600-inducing manipulations is that for pronouns the referent is highly active (Burkhardt, 2007; Kuperberg, 2007; Brouwer et al., 2012). A smaller number of studies have identified an Nref effect when a pronoun refers to a non-salient (or wholly absent) character, or when the gender of the pronoun causes the processor to chose a less active antecedent.

We hypothesized that pronouns with an unexpected referent that do not bear an accent are pragmatically infelicitous compared to accented ones. So in this case, accenting may reduce processing difficulty. It is not clear, however, what form this reduction of processing difficulty will take. Before going into that, it is necessary to briefly discuss relevant ERP research on accenting.

4.1.2 *The Effect of Accent on Comprehension*

There has been no research specifically investigating the unexpectedness accent in comprehension using ERP. We suggested that, based on the production work, the accent may help the comprehender deal with an unexpected concept, but how would "helping" manifest itself in the ERP signal, and is there any evidence for such an effect? Below we review previous research on pitch accent that is relevant to the present study: Accent has been found to induce positivities in the 100 to 350 ms time window. There is also some research examining the accenting of expected vs. unexpected content words, which may also be rel-

evant to our experiment. Finally, we examine the incorrect use of accent, looking at superfluous and missing accents.

First, there are clear changes in the ERP signal resulting from the physical, acoustic effect of accent as with any stimulus with stronger acoustic properties. For example, the processing of tones that are presented less often, have a higher F_0, and are louder leads to a greater intensity of the P200 (Carrillo-de-la-Peña et al., 1999). Because of this, when Heim and Alter (2006) found an early positivity in response to the acoustic properties of accent in the 250 to 350 ms time window, they described it as a P200, despite the latency differences. The literature has been consistent in finding a positivity around 200 ms for accented vs. unaccented stimuli (Dimitrova et al., 2012; Heim and Alter, 2006; Friedrich et al., 2001). However, the accent positivity might also represent a P300 effect, a component that results from stimulus frequency, and use of the stimulus to update context (Polich, 2007). Accented words occur less frequently than do unaccented words (in the broader context of language use, usually only one or two words are accented per sentence), and we argue here that they are used to indicate constituents that need to be processed more intensively. The P300 interpretation has also been discussed by Dimitrova (2012) and was not totally dismissed by Heim and Alter (2006) in the first place.

Turning to how the unexpectedness accent might affect comprehension of a pronoun that refers to an unexpected referent, the evidence suggests that the physical effects of accent may lead to deeper linguistic processing. Several studies have shown that accenting can give rise to a P600 effect and intensify N400 effects. Dimitrova et al. (2012) describe what they term an *accent positivity* which starts out like the P300 and then continues into later time-windows. They suggest that the latter half of the component involves discourse-related P600 activity of the type discussed by Brouwer et al., (2012, Burkhardt, 2007; Kuperberg, 2007).

Wang et al. (2011) found that the N400 amplitude increased for an anomalous word whether the constituent was accented or unaccented. In a dialogue about the preparation of a meal, one interlocutor asks what kind of *vegetable* was purchased. The answer specified either the congruent *eggplant* or the incongruent *beef*. Wang et al. found that in the accented condition there was

a greater N400 effect than in the unaccented analogue, suggesting that accent increases the intensity of processing, which is measured by the N400. Although the source of the unexpectedness is different in our experiment, we might expect a similar increase in the effect elicited by unexpectedness. However, in Wang et al., the N400 was associated with a category violation, and not with givenness, or world knowledge in general, as in the current experiment. Additionally, the lack of N400 effects on pronouns that we discussed above raises the question of whether an N400 will really occur.

Summarizing, the physical effect of accenting is predicted to cause a positivity in the 250 to 350 ms range (Heim and Alter, 2006; Dimitrova et al., 2012). There is also evidence that unlike a simple P200/P300, accent may cause a prolonged positivity, though Dimitrova et al. (2012) suggest the latter part of this positivity is a P600 effect associated with discourse-related processing. It is also possible that accenting a pronoun will increase an N400 effect through intensified processing, but given the general absence of N400 effects with pronouns this seems relatively unlikely.

4.2 THE PRESENT STUDY

The present study investigates the comprehension of unambiguous pronouns in spoken dialogues. If a pronoun that appears in a non-initial utterance of a dialogue refers to an unexpected entity, it may give rise to a P300 effect, an N400 effect, a P600 effect, an Nref effect or a combination of the four. Accenting this pronoun when its referent is unexpected is hypothesized to aid processing, and affect either one or all of the ERP effects mentioned above.

4.2.1 *Methods*

4.2.1.1 *Participants*

Twenty-one native speakers of Dutch were drawn from the University of Groningen community (10 men, 11 women, mean age 22.62). Prior to participating in the experiment they performed handedness inventories, and completed a questionnaire

to ensure they had no previous history of neurological problems, hearing problems, or dyslexia. Participants read and signed a consent form according to the Declaration of Helsinki before starting the experiment session.

4.2.1.2 *Stimuli*

In the present experiment, we examine the comprehension of accented pronouns in naturalistic, contextualized mini-dialogues (Table 4.1). The stimuli were manipulated for the expectedness of the referent, henceforth Expectedness (*expected*, *unexpected*), where we manipulated the pragmatics of the first part of the utterance to make the referent either expected or unexpected, and we manipulated the accent of the pronoun, Accent (*accented*, *unaccented*), so that the pronoun was accented or unaccented. See Table 4.1 for sample materials. English translations precede the original Dutch materials (in italics).

We created four versions of 120 spoken mini-dialogues following the same scheme as the item shown in Table 4.1. In the dialogues, the introductory utterance presented a context that invoked a common schema (*Sarah played terribly well and won the finals of the chess tournament...*). This was followed by a backchannel (confirmation of comprehension) utterance *oh?* or *mhm.* An alternative version was also created that led to a schema with quite different expected outcomes (*Sarah played terribly bad and lost the finals of the chess tournament...*). Lastly we presented a critical sentence, composed of a noun phrase, a transitive verb closely related to the context created in the first sentence, followed by an accented or unaccented pronoun for which there was only one referent explicitly available in the context (*The teacher congratulated her*). In order to preserve the naturalness of the utterance, we refrained from cross-splicing at the pronoun. Similarly, we chose not to add more lexical material after the pronouns, since this would have crucially changed the information structure. Depending on which version of the context sentence was used, the introductory utterance and the final utterance made the referent of the pronoun either expected or unexpected. Thus, utterances in the *expected* referent condition were designed to be plausible and to fit within a common schema of events; consistent with expectations based on world

Table 4.1: Sample of materials used in the experiment.

EXPECTED REFERENT UNACCENTED PRONOUN	EXPECTED REFERENT ACCENTED PRONOUN
A: Sarah played terribly well and won the finals of the chess tournament.	A: Sarah played terribly well and won the finals of the chess tournament.
B: Oh?	B: Oh?
A: The teacher congratulated her.	#A: The teacher congratulated HER.
A: *Sarah speelde vreselijk goed en won de finale van de schaakwedstrijd.*	A: *Sarah speelde vreselijk goed en won de finale van de schaakwedstrijd.*
B: *Oh?*	B: *Oh?*
A: *De juf feliciteerde haar.*	#A: *De juf feliciteerde* HAAR.
UNEXPECTED REFERENT UNACCENTED PRONOUN	**UNEXPECTED REFERENT ACCENTED PRONOUN**
A: Sarah played terribly bad and lost the finals of the chess tournament.	A: Sarah played terribly bad and lost the finals of the chess tournament.
B: Oh?	B: Oh?
#A: The teacher congratulated her.	A: The teacher congratulated HER.
A: *Sarah speelde vreselijk slecht en verloor de finale van de schaakwedstrijd.*	A: *Sarah speelde vreselijk slecht en verloor de finale van de schaakwedstrijd.*
B: *Oh?*	B: *Oh?*
#A: *De juf feliciteerde haar.*	A: *De juf feliciteerde* HAAR.

knowledge. Utterances in the *unexpected* condition were designed to be inconsistent with common schemas/expectations based on world knowledge.

These items were put through a process of pretesting to ensure that the expected and unexpected referents were sufficiently plausible or implausible. These pretests were given to a group of participants drawn from the same population as the experiment, but who did not actually participate in the experiment. The dialogues were presented without the backchannel response, in written form, in a two-condition randomized latin-square design. Participants rated how well the first utterance of the dialogue fit with last utterance of the dialogue. Unexpected stimuli were required to have less good fit than expected stimuli. We rejected utterances that did not meet our selection requirements, replaced them and re-tested the materials using the same procedure, which in practice took three rounds of pretesting (see the column utterance fit, overall row in Table 4.2).

After recording the sentences (see Acoustic Analysis section), we created item groups by matching the stimuli from one group to another, in order to eliminate the possibility that inequalities across groups that could influence our findings. We matched length of the first utterance (in ms), the fundamental frequency (F_0) of the verb in the last utterance, the degree of utterance fit (Table 4.2). T-tests performed between these groups found p's > .50 for utterance length and p's > .48 for verb frequency.

Versions of the items were distributed across four lists so that participants heard an equal number of items from each condition (all taken from one item group). The order was pseudo-randomized and participants never heard more than one version of any item.

ACOUSTIC ANALYSIS. The dialogues were recorded in a sound booth using Adobe Audition CS5. We recorded two native speakers of Dutch (one man and one woman) reading the felicitous versions of the dialogues, that is, containing the unaccented pronoun with the expected referent, and the accented pronouns with the unexpected referent. The speakers were asked to speak clearly and naturally. Dialogues were analyzed and cross-spliced using Praat 4.0 (Boersma and Weenink, 2009). We cross-spliced all dialogues at the pause after *oh?* and created the infelicitous

accented pronouns with an expected referent, and unaccented pronouns with unexpected referents. The sentences were re-recorded as necessary to control for clarity of speech, naturalness and to make sure the accent in the final sentence was produced clearly.

To ensure consistency of the prosodic pitch accent, we analyzed the prosody in the Verb and Pronoun windows. Figures 4.1 and 4.2 show the fundamental frequency and length for the accented and unaccented pronoun conditions. We first discuss the verb window, and then the pronoun window.

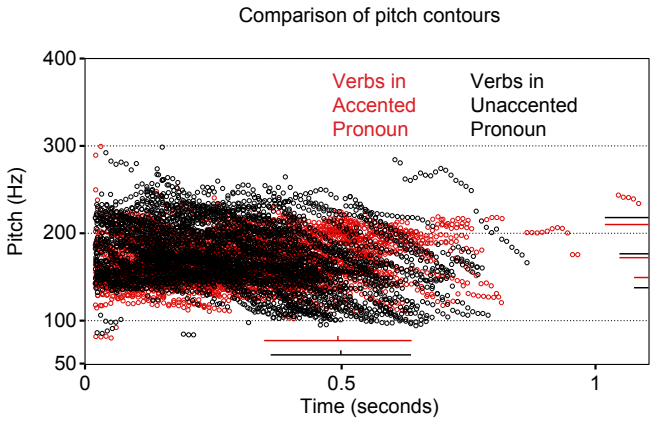

Figure 4.1: Pitch (F_0) and length of all verbs in the experiment. The horizontal lines above the Time axis indicates the average minimum length (left endpoint of horizontal line, average mean length (tick mark on the line), and average maximum length (right endpoint part of line). On the rightmost edge of the figure are three tickmarks. The longest represents the average maximum, followed in size by average mean and average minimum.

Figure 4.1 suggests that there was a difference between the two conditions at the Verb position. This difference was also evident in the statistics. We performed paired t-tests on the two conditions (Accented, Unaccented) to examine differences in duration and fundamental frequency. At the Verb, there was a significant difference between the two conditions in terms of Minimum F_0 ($p < .0001$) and Mean F_0 ($p < .05$). While in both

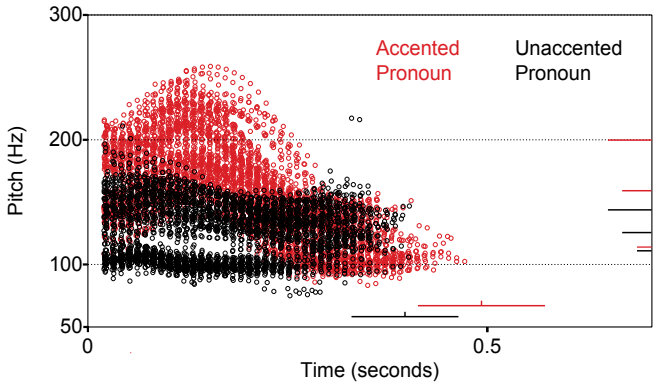

Figure 4.2: Pitch (F_0) and length of all pronouns in the experiment. The horizontal lines above the Time axis indicates the average minimum length (left endpoint of horizontal line, average mean (tick mark on the line), and average maximum length (right endpoint part of line). On the rightmost edge of the figure are three tickmarks. The longest represents the average maximum, followed in size by average mean and average minimum.

conditions, F_0 appears to start at the same level, the unaccented pronoun condition goes on to reach a higher maximum pitch and a higher mean pitch overall. The F_0 lowers toward the end of the sentence, producing a lower average minimum F_0. This suggests that our speakers hit a higher maximum point, and finished the word proceeding on a downward trajectory. In the accented pronoun condition, speakers again start at roughly the same point, but appear to move in a gradual upward trajectory toward the high accent on the next word, the pronoun. Verb statistics are given in numerical form in Tables 4.3 and 4.4.

In the pronoun window, the contrast between the two conditions is clear, both visually and statistically. In Figure 4.2, the pitch of the accented pronoun attains an obvious peak and descends, whereas in the unaccented condition the pronoun is pronounced with a flat F_0 the pronoun condition. There were significant differences between the two conditions: Accented pronouns were longer in duration, had higher Minimum F_0 ($p < .04$), higher mean Maximum F_0 ($p < .0001$), and a higher Mean F_0 overall ($p < .0001$). The slight gap visible within the two conditions is due to speaker variation, and is more readily apparent in the pronoun condition because there was much less variabil-

ity in length and pronunciation relative to the verb condition. See Table 4.4 for numerical values of pronoun F_0 statistics.

Qualitatively, in the accented condition, the speakers used an H*L (first rising, then falling) accent on the pronoun (following Transcription of Dutch Intonation guidelines), and they used a slight H* (rising) accent on the verb in the unaccented condition.

4.2.2 *Procedure*

When participants first entered the lab, they were asked to read the instructions, while being prepared for testing. They were then asked to sit in a comfortable chair in the testing room, with the monitor adjusted to a comfortable height. The computer volume was checked to make sure that the audio files were played at a comfortable volume. Participants were then given the instructions once again on the computer screen.

Each trial proceeded as follows: The beginning of a trial was indicated by the presentation of a cross that participants were told to fixate on. After 1000 ms, the introductory utterance and the backchannel response were played, followed by the critical sound file after 500 ms. It was determined ahead of time that this was a natural sounding interval. After the critical utterance, participants fixated on the cross for 1300 ms. After roughly one in ten items, participants saw a word. They were asked to indicate whether the word bore some relation to the immediately preceding mini-dialogue using either the 1 key (yes) or the 2 key (no) with their right hands. These were mostly abstract words. For example for the story about *Sarah* seen above, the word would be something like *competition*. The related words had an unpredictable but nevertheless evident relation to the story, whereas unrelated words were clearly unrelated to the story (for the story above, an example would be *mathematics*). Participants were told that if they simply concentrated hard on listening to content of the mini-dialogues, they would easily be able to answer the relatedness question. In practice, this allowed us to distract participants' attention from the main manipulation of the experiment, without the use of fillers. When asked, participants in fact thought the experiment was about the rela-

Table 4.2: Item characteristics.

		LENGTH OF BEGINNING (MS)		UTTERANCE FIT (ANTECEDENT COHERENCE)		
ITEM GROUP		EXPECTED	UNEXPECTED	EXPECTED	UNEXPECTED	INL VERB FREQUENCY
1	Mean	6735	6481	5.85	2.70	1.24
	Std dev	1038	898	0.64	0.79	0.50
2	Mean	6862	6628	5.55	2.47	1.22
	Std dev	1167	817	1.03	0.60	0.57
3	Mean	6704	6620	5.51	2.55	1.23
	Std dev	1017	1064	1.06	0.90	0.67
4	Mean	6828	6581	5.57	2.61	1.30
	Std dev	928	1031	0.61	0.60	0.64
Overall	Mean	6782	6578	5.62	2.58	1.25
	Std dev	1030	948	0.86	0.73	0.59

Note. INL = *Instituut voor Nederlandse Lexicologie* Dutch institute of lexicology.

Table 4.3: Duration of segments in critical utterances (ms).

ITEM GROUP		ACCENTED			UNACCENTED		
		FIRST WORD	VERB	PRONOUN	FIRST WORD	VERB	PRONOUN
1	Mean	496	474	486	516	492	393
	Stdev	161	121	57	160	131	67
2	Mean	473	536	479	488	532	401
	Stdev	158	174	73	125	122	61
3	Mean	498	496	505	506	515	388
	Stdev	158	150	101	193	159	63
4	Mean	509	465	500	511	455	406
	Stdev	196	119	82	191	126	77
Overall	Mean	494	493	493	505	498	397
	Stdev	167	144	80	167	137	67

Note. INL = *Instituut voor Nederlandse Lexicologie* (Dutch institute of lexicology).

Table 4.4: Fundamental frequency characteristics for verbs and pronouns (F_0 in Hz).

		VERB			PRONOUN		
		MIN	MAX	MEAN	MIN	MAX	MEAN
Accented	Mean	144.54	186.66	158.37	98.89	181.86	144.01
Pronoun	Stdev	22.15	75.45	24.78	18.34	27.21	21.99
Unaccented	Mean	107.19	177.93	143.46	96.38	127.71	108.51
Pronoun	Stdev	30.39	65.92	31.54	21.09	24.19	20.47

tion of the words, rather than about pronouns. Following each trial there was a 3000 ms rest indicated by stars (*****) during which participants could blink.

Participants took part in a practice session, which included 5 items. They were asked during the practice session if they could hear the sentences properly. After the practice session, they were asked if the procedure was clear to them, and if they had any questions. Participants were made to take three rest periods.

EEG RECORDING. Data were recorded using a 64–channel, extended 10–20 system cap with Ag/AgCl electrodes. A DC amplifier was used with a digital 67.5 Hz cutoff FIR filter to prevent aliasing. The amplifier sampled at a frequency of 250 Hz, amplifying online against the average of all scalp electrodes.

DATA ANALYSIS. The data were processed using Brainvision Analyzer 2.0 software. Re-referencing took place off-line to the average of the two mastoid electrodes. A band-pass filter of 0.0100 to 40 Hz, 48 dB/oct was applied, with a time constant of 15.9155s. Blinks were identified using bipolar horizontal and vertical electro-oculograms, and were corrected using the Gratton-Coles algorithm. An automatic rejection procedure examined and removed any data segments that violated any of the following requirements: a) a maximum voltage step greater than 20 μV, (b) maximum allowed absolute difference greater than 100 μV within a 100 ms moving window (c) amplitudes less than or greater than 150 μV relative to the initial value, or (d) activity less than 0.01 μV. Stimuli were time-locked exactly at the

Table 4.5: Electrodes pooled for each ROI.

	LEFT	MIDLINE	RIGHT
Anterior	FP1, F7, F3	Fpz, AFz, Fz	Fp2, F8, F4
Central	FC1, FC5, CP3	FCz, Cz, CPz	FC2, FC6, CP4
Posterior	P3, O1, P9	Pz, POz, Oz	P4, O2, P10

pronoun using the measurements made as part of the acoustic analysis.[1]

The inclusion criterion for participants was a minimum of 85% of the original data in the ROIs. Overall, this rejection procedure led to a 94.3% data inclusion rate, with even data retention across conditions and manipulations.

We chose 9 ROIs composed of 3 pooled electrodes each, aiming for the best possible coverage of the scalp distribution (Table 4.5). In the ANOVA analyses presented below we included Laterality (*Left*, *Midline*, *Right*) and Anteriority (*Anterior*, *Central*, *Posterior*) in our analyses.

The overall ANOVAs for each region were thus Laterality x Anteriority x Expectedness x Accent. Effects of Laterality and Anteriority are only discussed if they interact with one of the factors of interest.

We chose four time windows for analysis, based on the literature: 1) an Early time window (Magne et al., 2005), 2) an *Accent Positivity* time window (Dimitrova et al., 2012; Heim and Alter, 2006), 3) an N400 time window, (Schumacher and Baumann, 2010; Dimitrova et al., 2012) and 4) a late time window Osterhout and Mobley (1995); Nieuwland and Van Berkum (2008). We took a long (1000 ms) pre-stimulus baseline for the analysis of the pronoun, to minimize any pre-existing effects.

1 An analysis of the verb region revealed a significant main effect of prosody on the verb in the N400 window ($F(1, 13) = 6.869$, $p = .021$), with trends in the early positivity region and in a 500 to 700 ms window, suggesting a positivity for accent (since in the unaccented pronoun condition the verb had an H* accent). No other interactions or effects were significant ($p > .1$)

4.2.3 *Results*

An overall analysis indicated interactions of Anteriority, Laterality, Expectedness, and Accent that were significant in all time windows except the early one; there we found a marginally significant interaction between Anteriority and Expectedness (Table 4.6).

To determine the origin of these four-way interactions, we performed pairwise comparisons between the individual conditions. First, we looked at the effect of Expectedness (unexpected — expected) separately in the condition where pronouns were unaccented, and in the condition where pronouns were accented. Secondly, we investigated the effect of Accent (accented — unaccented) in the condition where the referent of the pronoun was expected, and in the condition where the referent was unexpected.

4.2.3.1 *Effects of Unexpectedness*

UNACCENTED PRONOUNS. Comparing unexpected with expected unaccented pronouns, as shown in Table 4.6 and Figures 4.3 and 4.4 in the *200 to 350 ms* range, we found a central positivity, evidenced by a trend across the Central ROIs taken together ($F(1, 13) = 3.208$, $p = .097$). This central positivity continued into the *350 to 500 ms* time window ($F(1, 13) = 3.77$, $p = .074$). This suggests that unexpected pronouns produced an early central positivity relative to expected pronouns. There was no effect in the late time window.

ACCENTED PRONOUNS. There were no significant differences until the *500 to 1200 ms* window, where the unexpected condition had a broadly distributed positivity relative to the Expected condition, with trends in right central ($F(1, 13) = 3.52$, $p = .08$), and left posterior regions ($F(1, 13) = 3.855$, $p = .08$), as visible in Figures 4.3 and 4.5.

4.2.3.2 *Effects of Accent*

PRONOUNS WITH EXPECTED REFERENTS. In the *200 to 350 ms* time window, accented pronouns had a central and slightly

Table 4.6: Probabilities and F values for significant effects and trends (< .1) by time window.

		EARLY WINDOW 100–200	ACCENT POSITIVITY WINDOW 200–350	N400 WINDOW 350–500	LATE TIME WINDOW 500–1200
Accent	F(1, 13)			3.276	
	p			0.09	
Expectedness	F(1, 13)				
	p				
Anteriority x Expectedness	F(2, 26)	2.974			
	p	0.07			
Anteriority x Laterality x	F(4, 52)		3.171	3.27	4.27
Expectedness x Accent	p		0.02	0.02	0.01

Note. Significant level of effects with p-level below .1

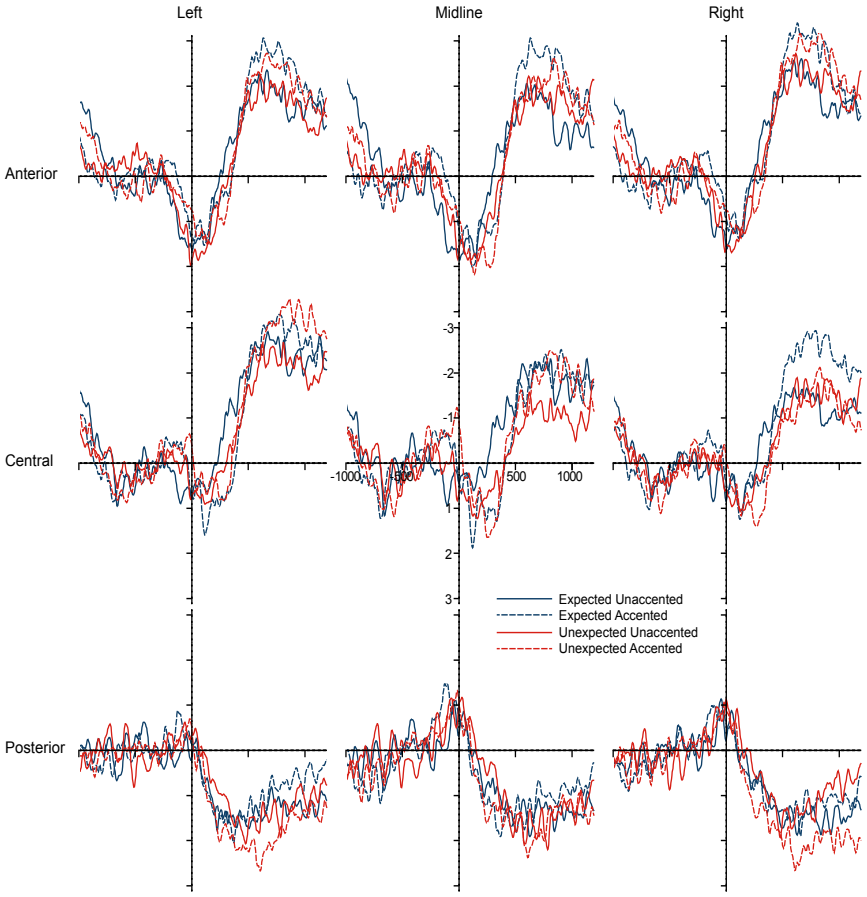

Figure 4.3: Grand average of ERP waveforms. Expected stimuli are plotted in dark blue, Unexpected in red, Accented with a continuous line, and Unaccented with a dotted line. In this and all following figures, negativity is plotted upward, the onset of the pronoun is at 0 ms.

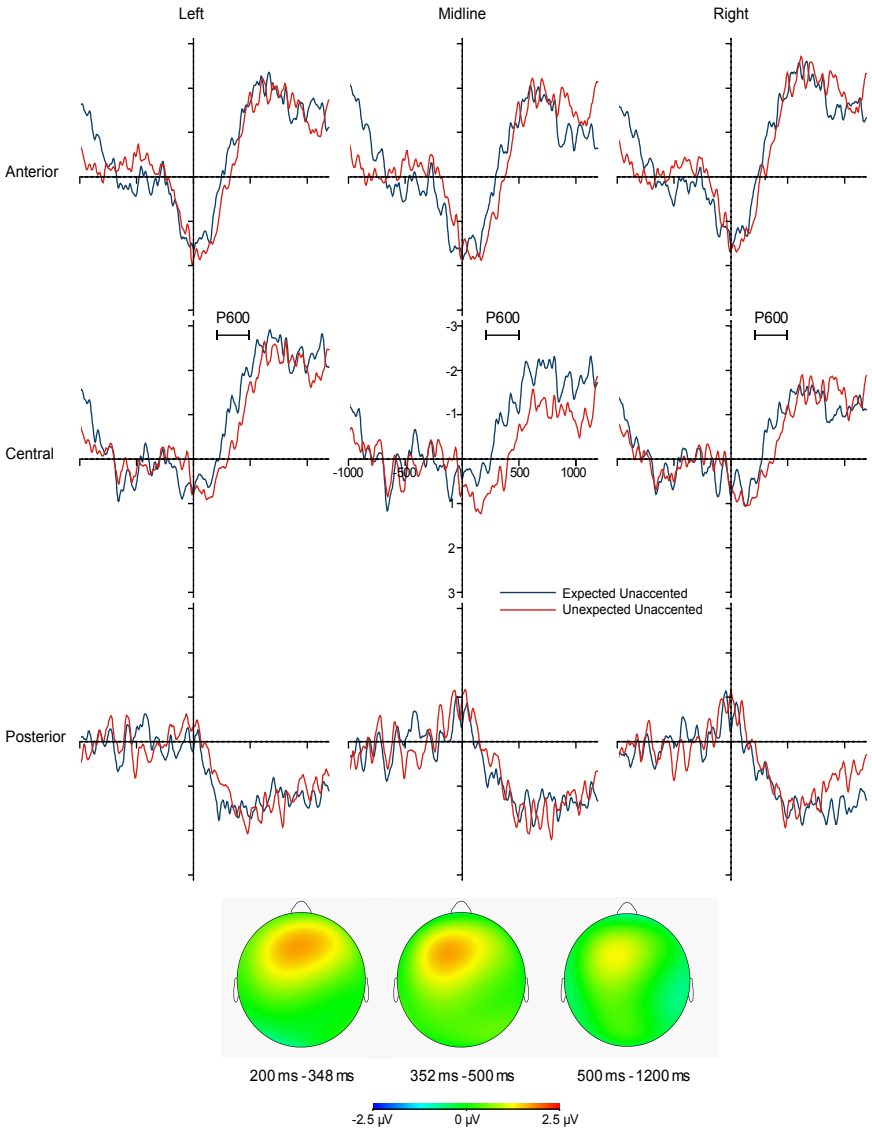

Figure 4.4: Effect of Unexpectedness for unaccented stimuli. Expected stimuli are plotted in blue, Unexpected in Red. Significant effects are indicated above the waveform, and the I-bar indicates the window(s) for which they were significant.

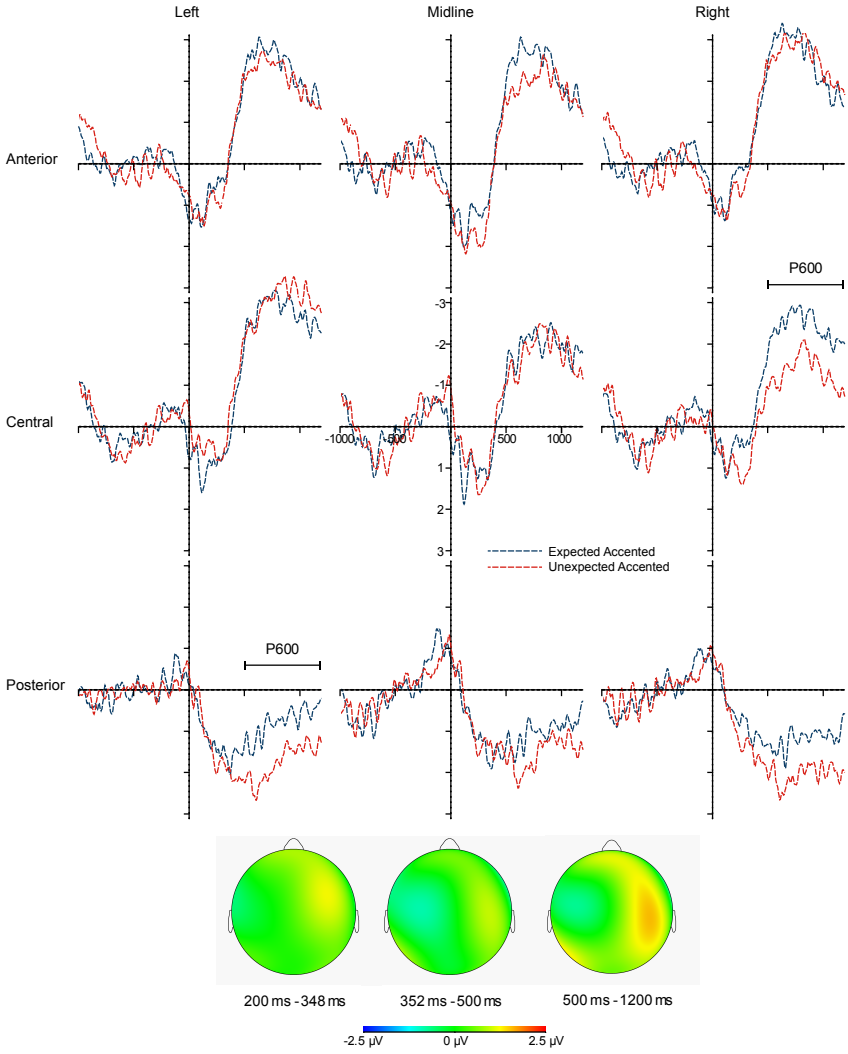

Figure 4.5: Effect of Unexpectedness for Accented stimuli. Expected stimuli are plotted in blue, Unexpected in Red. Significant effects are indicated above the waveform, and the I-bar indicates the window(s) for which they were significant.

left lateralized positivity, relative to unaccented stimuli, evidenced by a significant difference in the Left Central region ($F(1, 13) = 5.44$, $p = .036$) and the Central region ($F(1, 13) = 4.68$, $p = .050$). In the *500 to 1200 ms* time window, the same comparison (accented minus unaccented) showed a frontal-medial and right-central negativity, although the trend emerged only in the right central region ($F(1, 13) = 3.912$, $p=.070$). Visual inspection of Figure 4.3 supports this interpretation.

PRONOUNS WITH UNEXPECTED REFERENTS. In the *200 to 350 ms* time window, comparing accented to unaccented pronouns within the unexpected condition there is a broad positivity spread across most of the scalp; taking the midline ROIs together (Mid-Anterior, Mid-Central and Mid Posterior), we found a significant effect of accent ($F(1, 13) = 5.188$, $p = .040$), with a trend in the right central region ($F(1, 13) = 4.192$, $p = .061$). In the *500 to 1200 ms* range, accented pronouns were more *negative* than unaccented pronouns, in the left Central region, $F(1, 13) = 6.766$, $p = .022$). The same comparison in the left posterior ROI indicated a trend towards a *positivity* $F(1, 13) = 4.28$, $p = .059$).

4.2.3.3 *Summary of Results*

Just referring to an unexpected antecedent provoked changes in processing early on, irrespective of whether the pronoun was accented or unaccented (Table 4.7). Accenting the pronoun did, however, change processing: it both intensified early processing and provoked additional late processing. Unexpected pronouns created a central positivity from 200 to 350 ms in both accented and unaccented conditions, although this was more right lateralized in the accented condition. Accented unexpected pronouns showed additional activity in the 500 to 1200 ms time window, a left central negativity and a left posterior positivity.

Looking at the data split into Expected and Unexpected conditions, it is clear that regardless of expectedness, accent created central positivities in the 200 to 350 ms time window, and central negativities in the 500 to 1200 ms time window. The early accent positivity was more left-lateralized for expected stimuli, and more right-lateralized for unexpected stimuli. In the 500 to 1200 ms time window, in the Unexpected condition, accent

Table 4.7: Comparisons by condition

COMPARISON	ACCENT POSITIVITY WINDOW 200–350	N400 WINDOW 350–500	LATE TIME WINDOW 500–1200
Unaccented (Unexpected—Expected)	+ Central ROIs	+ Central ROIs	
Accented (Unexpected—Expected)			+ RCentral +LeftPost
Expected (Accented—Unaccented)	+LCentral* +MidCentral		—Rcentral
Unexpected (Accented—Unaccented)	+Midline ROIs* +RCentral		—LCentral* +LPosterior

Note: All effects below at least p = .10, effects below p = .05 indicated with "*". +/— indicates the direction of the difference.

was associated with a left posterior positivity in addition to the central negativity. The central negativity in the late time window associated with accent differed in lateralization by expectedness. In the Unexpected condition it was left central, and in the Expected condition it was right central. This suggests that a central negativity was decreased by the central positivity from the earlier time windows, and that both conditions in fact show a central negativity in interaction with a central positivity.

To summarize, Accent and Expectedness interacted with each other throughout the entire epoch. In the earliest time window (100–200) there was a strong tendency toward a main effect of Accent, replicating earlier studies, but in the accent positivity window (200–350), the positivity associated with accent was modified by expectedness—only unaccented pronouns showed an effect of Expectedness. This pattern continued, less distinctly, into the N400 time window (300–500). In the late time window, accent elicited a negative response, which continued to interact with Expectedness, and seems to result from only accented pronouns showing an Expectedness effect in this time window.

4.3 GENERAL DISCUSSION

Previous research (Dimitrova et al., 2009) has suggested that speakers can use accent to draw attention to unexpected constituents. Assuming that listeners make use of cues supplied by speakers, we hypothesized that accenting a pronoun that refers to an unexpected referent would aid comprehension. In this experiment we investigated how accent influences pronoun interpetation, and varied whether the pronoun referred to a character that was expected (based on previous discourse and world knowledge), or unexpected. In addition to taking a first look at the unexpectedness accent using ERPs, we also examine the processing of pronouns from a new perspective: Past studies have either looked at pronouns with absent or ungrammatical referents (Osterhout and Mobley, 1995) or have investigated how an expectation based on grammatical structure and pragmatics affects resolution of the pronoun to a non-preferred antecedent. In the present experiment, we look at a "pure anomaly" case where only one antecedent is available but it is unexpected.

We found that Expectedness and Accent interact in three different time-windows. For unaccented pronouns only, in the earliest time-window (200–350 ms) and in the later N400 time-window (350–500 ms), we found a positivity for pronouns with an unexpected referent, as compared to pronouns which are congruent with the context. There was no evidence of an N400 effect. The positivity was only present in unaccented pronouns; it was absent when pronouns were accented. In the following late time-window (500-1200 ms), accented pronouns with unexpected referents instigated a long-lasting positivity as compared to controls; no such effect was present for unaccented pronouns. We first examine these results for each level of expectedness separately, and later discuss effects at each level of accent (4.3).

4.3.1 *Unaccented*

4.3.1.1 *Accent positivity window*

There is evidence that pragmatic processing occurred at a very early stage, as shown by the positivity for unexpected referents relative to expected referents, starting in the 200 to 350

ms time window. The fact that the expectedness effect is seen for *unaccented* elements runs contrary to one of the possible outcomes discussed in the introduction. There, we speculated that accenting could lead to increased processing, as suggested by Wang et al. (2011), but these results suggest that also the unexpected element is detected quickly. Despite the relatively early latency, we interpet this as an early P600 component similar to effects found for pragmatic violations (Lau et al., 2006; Hoeks et al., n.d.). This is compatible with accounts suggesting that the P600—early or late—results from processes involved in integrating a word into a conceptual model or reanalysis of an existing model (Burkhardt, 2007; Brouwer et al., 2012). For instance, Osterhout and Mobley (1995) found a P600 component when the gender of the antecedent did not match the available referent in a subordinate clause, a finding replicated by Nieuwland and Van Berkum (2006). There have also been a number of P600s reported for violations of an initial bias, for example an implicit causality bias (Van Berkum et al., 2007), or biases that verbs induce for likely continuations (Ferretti et al., 2009; Koornneef and Sanders, 2012).

There is a common thread running through our study, and other studies that have found early P600s, like Hoeks et al. (nd) and Lau et al. (2006). In these studies expectations were highly constrained before the onset of the critical word. Lau and colleagues found that violation of an expectation based on word order created a positivity as early as 200 ms. There have also been early P600s in response to strong pragmatic biases. Hoeks et al. found early P600s in response to pragmatic violations occurring in response to question answer pairs. In the pragmatic-violation condition, a question was asked about two characters, but (in Dutch) the presence of an adverb indicates that the second noun must be interpreted as a direct object rather than the beginning of a second clause and thus that the response only addressed half of the question (Q: *What did the mayor and the alderman do?* A: *The mayor praised the councillor and the alderman* exuberantly.)

Our experiment is much like these previous studies, since the pronoun itself adds very little semantic or pragmatic information to the expectation, other than confirming or contradicting it. In the expected condition, the discourse leading up to the

pronoun made the only available referent of the pronoun a very plausible continuation. In the unexpected condition, the discourse made the only available referent very implausible. In both conditions, these biases occurred before the pronoun. By the time a comprehender reached the pronoun, they had a great deal of pragmatic information at their disposal, biasing them either to or away from a single available antecedent.

To explain the early occurrence of the P600, we favour the strong fit hypothesis discussed by Lau et al. (2006). They suggest that when there is very strong evidence pointing toward a particular continuation, P600s occur earlier than when the expectation is weaker. One reason why there is a relative dearth of early P600s is likely to be that expectations typically would elicit an overlapping N400 component. If a positivity occurs in the same time window as a negativity, it will be cancelled out. In our study, this was not a factor, because we found no evidence of an N400 component.

4.3.1.2 *Unexpectedness and the N400*

Previous research has found N400 effects for pragmatically unexpected words; in particular for low-cloze probability words relative to high-cloze probability, and for words in an associated context relative to words in an unassociated context (Kutas and Federmeier, 2011). At the discourse level, N400s have been found for anomalous words relative to coherent words (Nieuwland and Van Berkum, 2006). We are not aware of previous research that manipulated the fit of a pronoun's referent using world knowledge. Previous research has examined initial biases to refer to a particular antecedent, and whether the pronoun actually refers to the character the bias points at, or the character the bias points away from, and neither character was implausible on the basis of the actual real world interpretation. This study manipulated interpetation given world knowledge

We set up a comparison between pronouns that referred to an expected referent and pronouns with an unexpected referent. The referent was expected or not, based on the scenario that was set up in a previous utterance, and pragmatic information introduced by the verb in the final sentence – for someone who *loses a tournament* to be *congratulated* is unexpected, but

for someone who *wins a tournament* to be *congratulated* is expected. There was no evidence of an increased N400 to unexpected antecedents relative to expected antecedents.

The lack of an N400 in our experiment, taken together with previous research on pronouns, suggests that there is something different about how pronouns are integrated into the context relative to content words. A possible explanation is that pronouns access information in the mental representation of the discourse that is built up and maintained in short term memory. Nouns, in contrast to pronouns, have a decontextualized meaning; when presented out of context, *dog* refers to a member of the category *dog*, and causes access to information about *dogs* in semantic memory. A pronoun like *him* takes nearly all of its meaning from discourse, and that information is by definition already active in semantic memory.

This is a feature that separates our study from past studies. Van Berkum et al. (1999), like us, manipulated the discourse context before a target word (e.g. *slow*) so that it fit or did not fit, and found an N400. This word added more information to the discourse, whereas the pronoun merely indicated that an already present character was part of an already-mentioned event in an unexpected way. *Slow* requires recourse to semantic memory, whereas *him* requires access to the discourse model.

This view of the difference between pronouns and lexical targets like *slow* is consistent with the view that the N400 primarily reflects the effort of activating information from semantic memory. This hypothesis is clearly articulated in the retrieval-integration account Brouwer et al. (2012), where words that are more difficult to *retrieve* from long-term memory create a greater N400, and where P600 amplitude reflects the *integration* of this information into the discourse model (although see Kuperberg et al., 2011). The early P600 in the current study is also consistent with this view. Pronoun resolution may involve accessing the current discourse representation in short term memory first, and then problem solving later if the search for a referent is unsuccessful.

4.3.2 *Accented*

4.3.2.1 *Late effect*

In the *unaccented* condition, when comparing pronouns with expected referents with those that have unexpected referents, we found evidence of an early P600. However, in the Accented condition we found evidence of a late difference. This later positivity for pronouns with an unexpected referent differs from the early P600 reported in the Unaccented condition not only in timing, but also in topography. The topography of the later P600 is more classic in that it was in the posterior part of the scalp (left posterior) although it also showed significant right central positivity. The early P600 was less classical, with a central topography, somewhat similar to Osterhout and Mobley (1995), who found central and posterior positivities. This suggests that these findings represent two distinct processes in comprehension. We hypothesized that accenting would help the listener in some way with integrating the unexpected information, which could for instance be manifested as a reduced positivity, that is, if we take the size of the P600 effect as indicative of processing difficulty. This is clearly not what we found. On the other hand, accenting *expected* information is not felicitous, and may give rise, especially where pronouns are concerned, to an increase in processing, due to the processor attempting to switch referents (cf., Taylor et al., 2013), which may in turn give rise to a *negativity* for expected referents (instead of a positivity for unexpected referents). In the current study, no actual alternative antecedent was available, but that does not necessarily ensure that our participants did not look for them. The negativity elicited by this search process may thus be seen as an Nref effect.

4.3.3 *Expected and Unexpected*

When comparing accented to unaccented items within the early time window, we replicated an early positivity, which has previously been shown with the same topography and the same time window in the literature. In the later time window, ac-

cented pronouns were associated with negativities that may be related to the Nref effect.

4.3.3.1 *Accent positivity*

In the present study, findings in the accent positivity time window replicate the positivity that was found in previous studies. There was a central positivity in response to accenting, both for the expected referent and the unexpected referent.

Accent thus seems to cause an early positivity, possibly resulting from both the physical properties of the stimulus, and some effect of the stimulus on linguistic processing. Most studies comparing accented words to unaccented words have found a positivity in time windows starting from about 100 ms post-stimulus onset and peaking at 250 to 350 ms (Dimitrova et al., 2012; Heim and Alter, 2006; Magne et al., 2005).

These positivities have been interpeted in several different ways: Heim and Alter, (2006; Friedrich et al., 2001) link accent-positivities to a P200, citing the increased physical energy of the stimulus (Carrillo-de-la-Peña et al., 1999). An alternative interpetation is that an accented constituent is more "surprising" than an unaccented constituent, resulting in a P300 (Magne et al., 2005). Dimitrova et al. (2012) point out that it may not in fact be a matter of P200 vs. P300, and that the positivity can in fact result from some combination of a P200 to the physical effects of the stimuli, and a P300 resulting from the use of accent to draw attention to the constituent.

4.3.3.2 *Nref*

Our results indicate that accenting leads to a central negativity for accented relative to unaccented stimuli for both expected and unexpected pronouns. Accenting a pronoun suggests that there is some violation of expectation. This expectation may be caused by reference to the most active antecedent, real world plausibility, and possibly even other things. We would like to assume that the negativity we found is related to referent search. The idea that accent triggers a broader search for an antecedent is reflected by the Nref, and is consistent with findings in the literature that accent leads to the choice of an initially dispreferred referent (Taylor et al., 2013; Kaiser, 2011; Venditti et al.,

2002). The presence of a negativity for both expected and unexpected pronouns suggests that accenting indeed causes the processor to look harder for an alternate referent whether the referent of the pronoun is plausible or not. The statistical interaction, however, suggests that the expected pronouns are associated with a more extensive search than pronouns with an unexpected referent.

4.4 CONCLUSION

In this experiment, we investigated the effects of accent and unexpectedness on pronoun resolution. Sentences that had a pronoun with an unexpected referent and described a situation that went against world knowledge (e.g. *someone loses a tournament and is congratulated*) did not produce an N400 effect, but an early P600 effect instead. This P600 is taken to reflect the increase in processing effort that is expended to create a coherent representation of the utterance. Accenting the pronoun most likely led to an Nref effect, reflecting the effort associated with trying to carry out a reference switch (but not succeeding). Alternatively, it is possible to interpret this effect as a reduced P600, perhaps indicating that interpretation did become easier by having an accent on the pronoun

5

CONCLUSION

This thesis looked at anaphora resolution, concentrating in particular on the influence of the markedness of the anaphoric expression (in terms of prosodic and/or morphological realization) and plausibility (based on world knowledge) of the antecedent. One of the key issues in this type of investigation is how diverse sources of information are combined by the language processor. I tried to answer this by using a diverse set of psycholinguistic response measures, such as self-paced reading, antecedent selection and ERP (Event-Related brain Potentials) registration. In this conclusion section, I discuss what my findings say about what a model of pronoun comprehension should look like, and what it should not look like in order to explain these and other results.

The investigation began by looking at sentence-initial pronouns which were either ambiguous or unambiguous in the discourse context. The ambiguous pronoun's reference was later resolved on the basis of world knowledge (Chapter 2). In the experiment, serial models were contrasted with parallel models, since these models make different predictions with respect to when and how information is used. Some past evidence from language processing has found support for parallel models, which predict that all information is used immediately, while other evidence has supported serial models, which predict that the use of some information is delayed to a later stage. I posited in conjunction with Swets et al. (2008) that differences between past findings can potentially be explained by task effects.

The specific serial processing model that was tested, the Unrestricted Race model, predicted that temporarily ambiguous pronouns would be faster than unambiguous pronouns at the initial point of ambiguity, but at a later point of disambiguation the

ambiguous condition would be slower than the disambiguated condition. The exemplar parallel model, the Constraint-Based processing model, predicts that ambiguity will cause slower processing. I included three task manipulations to check whether task effects were indeed a plausible explanation for differences in earlier results. Participants either 1) indicated who they thought the antecedent was after reading the story, 2) rated the story for naturalness, or 3) did not have a secondary task.

The results suggested that the Constraint-Based Model best described processing in the cases where participants made a particular effort to understand the stories, since there was evidence that ambiguous pronouns take longer to comprehend than unambiguous pronouns in the Disambiguation Question and the Plausibility Rating conditions. For both of these tasks, pronoun interpretation was crucial. There was no difference between ambiguous and unambiguous pronouns in the task condition that required mere self-paced reading. These results combined suggest that the Constraint-Based model should be modified in such a way that it can incorporate some kind of *task-engagement* parameter. One mechanism that could explain the results is a monitoring mechanism that continuously checks the outcome of the constraint-satisfaction process, as some implemented Optimality accounts do (Van Rij et al., 2013). If an answer is required after reading the story, as in the Antecedent Selection task, the monitor will not be content until the processor has found clear evidence for either of the possible antecedents, leading to prolonged competition. If, on the other hand, no answer is required, monitoring may remain superficial, rendering competition unnecessary.

Although I found support for the Constraint-Based model in referential ambiguity processing, there has not been massive support for the Constraint-based model in other fields of ambiguity resolution (e.g. syntactic ambiguity resolution). It is possible that this has to do with the specific research paradigm that has been used in the majority of studies looking at parsing ambiguities, namely self-paced reading. As it happens, support for the Constraint-based model (not specific to reference comprehension) has come from other methodologies: Eye-tracking research using the visual world paradigm has tended to find evidence for the Contraint-Based model of ambiguity resolution

(Spivey et al., 2002), as has mouse-tracking (Farmer et al., 2010). The evidence from reference comprehension seems to favour a parallel model where ambiguity leads to longer processing times. This is in agreement with ERP-studies of reference comprehension, which found that ambiguous pronouns gave rise to a negative shift in the ERP waveform as compared to unambiguous stimuli (NRef; Van Berkum et al.).

A Constraint-Based processor normally makes use of a number of factors in deciding on the best alternative. In the experiments described in Chapter 3, I looked at the complex interaction between world knowledge, markedness, order of mention, and syntax when selecting an antecedent. Markedness and world knowledge were shown to outweigh syntactic and order-of-mention cues in the materials used here. World knowledge was capable of dictating the referent to a strong degree. The results were different when pronouns were equiplausible, however, i.e., when there was no (world knowledge) bias towards either of two possible referents. For equiplausible pronouns, marked delivery of the pronoun, either by accentuation or by adding morphological complexity, was able to switch the referent. When there was a strong bias toward one antecedent, however, a marked pronoun reduced the preference for that antecedent, but did not switch it.

These results suggest that factors that determine pronoun resolution must be able to interact with one another, irrespective of whether the factor is syntactic or pragmatic. This is compatible with the standard Constraint-Based model that were described in this thesis, but also with other constraint-based theories such as Optimality Theory, that describes decision making processing on the basis of ordered constraints in (off-line) semantic interpretation (e.g. Hendriks and de Hoop, 2001) and on-line sentence comprehension (Hoeks and Hendriks, 2011; Lamers and de Hoop, 2005).

In Chapter 4, I examined the neural processes underlying pronoun comprehension. The study focused on the interaction of markedness and world knowledge in pronoun comprehension. For the experiment in Chapter 4, I measured ERPs while participants listened to sentences where the only available referent was either expected as the referent of the pronoun (e.g. someone won a tournament and was congratulated) or unexpected

(e.g. someone lost a tournament and was congratulated). Past research by Dimitrova et al. (2009) found that expressions referring to unexpected constituents were more likely to receive an accent. One possible interpretation of this fact was that accenting could perhaps make comprehension of an unexpected referent easier to understand. On the other hand, it might very well be the case that speakers were merely acting on a reflex to accent something that was a surprise for them. In that case, listeners would perhaps not be helped at all.

The results from the ERP study in Chapter 4 leave some room for interpretation and follow-up. In pronouns that did not carry an accent, there was an effect of unexpectedness (as compared to pronouns with an expected referent), interpreted as a P600 effect. Accenting the pronoun did not eliminate or even reduce this positivity, although with prosodic focus, the positivity started much later than in unaccented pronouns. This raised the question of whether I am looking at the same component, but shifted in time, or if this positivity should perhaps receive a different interpretation. For instance, it is entirely possible that this is not a positivity for pronouns with unexpected referents, but rather a negativity for accented pronouns with expected referents. Indeed, it seems plausible that accenting a pronoun without any apparent reason must induce extra processing, most likely the search for an alternative referent. The search for an alternate referent relates to the switch reference effect found in the antecedent selection studies reported in Chapter 3. This extra processing may thus have taken the form of an Nref effect (Van Berkum et al., 2007) an ERP-effect associated with a search for potential referents or disambiguation between referents (Nieuwland et al., 2007). This suggests that accenting provoked more intensive processing, but only in the case of the pronoun with the expected referent.

It is unclear whether accenting the pronoun with the unexpected referent completely reduced processing difficulty; I must leave it to future research to disentangle whether there is indeed an Nref for expected referents from the possible positivity for unexpected referents. Finding evidence for such a positivity would not be very surprising. An efficient swimming stroke is not one where no effort is expended. It is one where force is exerted in just the right place to move the swimmer forward.

"Good" accentuation, like a good swimming stroke will cause comprehenders to expend their effort at just the right place.

The ERP study suggests that use of a more marked form, like an accented pronoun, may lead to an intensified search for antecedents. In Chapter 3 I found that a marked anaphor was strong enough to reduce preference for an initially highly preferred antecedent. I also saw that when there was an equal amount of bias to two potential referents, marking the pronoun changed the preference to the alternate antecedent. The ERP results suggest that hearing an accent caused participants to engage in a search for an (unavailable) alternate antecedent. Future research should look to see whether marked constituents in general trigger similar ERP components.

5.1 FUTURE DIRECTIONS

Understanding reference is a prerequisite for successful communication. Pronouns and other referential expressions are particularly interesting as objects of study, since their processing involves many different factors that can potentially interact during comprehension and production. In languages like English and Spanish (Chapter 3) it appears that knowledge of the world outweighs markedness, at least in our stimuli, but possibly more generally. Whether this is true for all languages and all situations needs to be investigated. These results hint at a typological universal, and future research should look for exceptions.

Furthermore, our findings are in full agreement with the Neo-Gricean perspective on pronouns, which suggests that the markedness of a referring expression should be treated as a pragmatic, communicative signal to engage in a non-default route of processing. Though this may be universal, the form markedness takes is of course different for different languages. In English, an accent makes the constituent more marked. However, the referring expression would also be more marked if breathy voice were used, or if other forms of emphasis were used. Future research must determine which linguistic (or paralinguistic) features can function as a vehicle for markedness. In addition, we could find out more about markedness if we look at how L1 speakers of one language deal with an L2 that uses a differ-

ent system. There is evidence that in pronoun interpretation speakers use the weighting of factors from their first language to interpret their second language (Morrett and MacWhinney, 2013), and this could very well hold true for markedness and world knowledge.

A

APPENDIX: CHAPTER 2 STIMULI

A.1 DESCRIPTION OF STIMULI

The following list is the stimuli used for the experiment in Chapter 2. In parentheses, we present the word used to change the referent of the pronoun from first mention to second mention. Gender disambiguating sentences were created by changing the gender of one of the names. If the sentence referred to the first-mentioned character, then the name of the second character would be changed. Thus for sentence 1, the first-mention unambiguous sentence would begin *Richard zag Charlotte ...*

A.2 STIMULI

1. *Richard zag Roland de appelboom snoeien en kwam helpen. Roland was blij dat Richard de ladder vast wilde houden. Hij verbaasde zich over de zwaarte van de ladder (tak) die rechtgehouden moest worden.*

2. *Peer was te bang om met Teun de Martinitoren te beklimmen. Teun liet Peer daarom beneden wachten. Hij zwaaide vanaf de grond (toren) naar zijn maatje.*

3. *David vroeg Luuk te helpen bij het repareren van de wasmachine. Luuk hielp David achter de machine te kruipen. Hij las staande achter (voor) de machine de gebruiksaanwijzing voor.*

4. *Remco stond met Tim boven een betoverd vloerkleed. Tim zag Remco door het vloerkleed heen gaan. Hij gebaarde van onder (boven) het kleed dat het gevaarlijk was.*

5. *Henri nodigde Klaas uit om naar zijn optreden te komen. Klaas kreeg kaartjes van Henri. Hij knipoogde vanaf het podium (vanuit de zaal) veelvuldig naar hem.*

6. *Kevin belde met een behoorlijkebloedneus aan bij Wim. Wim deed nietsvermoedend de deur open voor Kevin. Hij gilde van het lachen (de schrik) toen hij het gezicht van zijn vriend zag.*

7. *Jelmer bood aan om Fokko rond te leiden in het Groninger Museum. Fokko belde Jelmer op om te zeggen dat er in Zwolle problemen met het spoor waren. Hij wachtte daarom op station Groningen (Zwolle) tot de vertraagde trein zou komen.*

8. *Gerrit stuurde Simon op een geheime missie naar een eilandje voor de Russische kust. Simon communiceerde in het geheim met Gerrit. Hij maakte gebruik van een radiozender om vanaf het vasteland (eiland) belangrijke informatie door te geven.*

9. *Sjoerd vroeg Freek om te helpen bij het renoveren van het dak. Freek hielp Sjoerd het dak op. Hij gaf de dakpannen vanaf het dak (vanaf de grond) aan zijn maatje.*

10. *Folkert vertelde Sjef dat hij op ruimtemissie zou gaan. Sjef omhelsde Folkert bij het afscheid. Hij keek later vanuit de spaceshuttle (vanaf het balkon) naar de sterren.*

11. *Dirk zong voor Bram een liedje in een karaoke-bar. Bram bedankte Dirk heel hartelijk. Hij begon vanaf het podium (vanuit het publiek) vrolijk te zwaaien.*

12. *Alex stuurde Bernhard naar Parijs voor een verslag van de finish van de Tour de France. Bernhard vertelde Alex dat de hele stad in een feeststemming was. Hij sprak vanuit de studio in Hilversum (Parijs) over de grote aantallen bezoekers.*

13. *Koen wilde aan Joop zijn auto wel uitlenen voor een ritje naar Parijs. Joop vroeg of Koen de auto 's ochtends al naar zijn flat kon brengen. Hij kreeg op weg naar de flat (Parijs) echter een klapband.*

14. *Gerard vroeg Floris naar de weg. Floris wees Gerard de juiste plek op de kaart. Hij vond het best wel moeilijk om de uitleg te snappen (geven) over de route.*

15. *Ruben gaf Jonas opdracht voor een foto-reportage over een zeldzame krokodil. Jonas vroeg Ruben om een speciale, supergevoelige camera. Hij wilde vanuit zijn functie als opdrachtgever (fotograaf) graag de beste resultaten.*

16. *Karel kreeg in de slaapkamer behoorlijk ruzie met Boris. Boris sloot Karel op in de badkamer. Hij schreeuwde vanuit de badkamer (slaapkamer) dat het afgelopen moest zijn met spelletjes.*

17. *Geert ging naar Johannes om een potje te schaken. Johannes versloeg Geert bij de eerste partij. Hij liep met een geïrriteerd (triomfantelijk) gezicht naar de keuken om een drankje te halen.*

18. *Aaron zei tegen Joris dat hij met een duikboot ging zoeken naar de reuzeninktvis Joris omhelsde Aaron bij het afscheid. Hij zond vanuit de duikboot (vanaf de wal) fascinerende informatie over het ondier.*

19. *Steven stuurde zijn zoon Ruud naar zomerkamp. Ruud vroeg Steven om te helpen met pakken. Hij schreef vanuit het ouderlijk huis (zomerkamp) een lieve brief.*

20. *Michiel vertelde Berry over problemen met de stoelgang.Berry gaf Michiel een uitermate werkzaam laxeermiddel. Hij riep vanuit het toilet (de woonkamer) dat het WC papier bijna op was.*

21. *Jules vroeg aan Sjaak of hij een sigaar mocht opsteken. Sjaak stuurde Jules meteen naar buiten, de tuin in. Hij zwaaide vanuit de tuin (woonkamer) vrolijk naar zijn vriend.*

22. *Herman vroeg Harrie om hem te helpen zijn klassieke auto te repareren. Harrie hielp Herman om onder de auto te komen. Hij zocht onder (naast) de auto naar een schroefje.*

23. *Jaap daagde Pim uit voor een kaartspel. Pim verspeelde al zijn geld aan Jaap. Hij verliet met een vrolijke grijns (droevige blik) de kamer.*

24. *Felix belde Hendrik op kantoor op om te zeggen dat hij ziek was. Hendrik wenste Felix een voorspoedig herstel. Hij ging snel weer naar de slaapkamer (het werk) terug.*

25. *Rik haalde een grap uit met Rob. Rob duwde Rik daarna in het meer. Hij begon vanuit het water (vanaf de kant) ontzettend hard te lachen.*

26. *Max bood aan Arie, die uit Antwerpen zou komen, rond te leiden in Amsterdam. Arie belde Max om te zeggen dat er vertraging was. Hij wachtte op Amsterdam (Antwerpen) Centraal op de vertraagdetrein.*

27. *Jacob was de advocaat van Sven in een geruchtmakende moordzaak. Sven overtuigde Jacob van zijn onschuld. Hij dacht op zijn advocatenkantoor (in zijn gevangeniscel) na over de komende rechtzaak.*

28. *Pieter waarschuwde Jasper voor de grote walvis die naar hun bootje zwom. Jasper werd met huid en haar opgeslokt, tot groot verdriet van Pieter. Hij zat dagen in de boot (walvis) voordat hij gered werd.*

29. *Wouter huurde Jeroen in om een portet te schilderen. Jeroen wist van Wouter dat het voor het nieuwe kantoor was. Hij ging daarna elke dag om te poseren (schilderen) naar het prachtige atelier.*

30. *Adam stuurde Toby op een missie naar het slinkende Amazone woud. Toby vroeg Adam om voortdurend in contact te blijven. Hij stuurde vanaf de thuisbasis (vanuit het oerwoud) zeer belangrijke berichten.*

31. *Leon streed met Otto om de eerste plaats in het klassement. Otto haalde in de laatste bocht Leon in. Hij reed daarna met een boos (blij) gezicht over de finish.*

32. *Thijs riep Roel om ook in de boot te komen. Roel zei tegen Thijs dat de boot teveel wiebelde. Hij giechelde vanaf de boot (oever) over de reactie van zijn vriend.*

B

APPENDIX: CHAPTER 3 STIMULI

B.1 ENGLISH EXPERIMENT 1A, EQUIBIASED

1. *Seth saluted Ronald with a toast, and Amy serenaded him.*
2. *Zachary seriously injured Luis, and Caroline bruised him.*
3. *Jeremy deceived Nicholas in a business deal, and Crystal misled him.*
4. *Carter savagely ridiculed Jeffrey, and Emma mocked him.*
5. *Trevor annoyed Carlos with probing questions, and Harriet irritated him.*
6. *Jack baited Manfred with crude remarks, and Lisa badgered him.*
7. *Luke disheartened Donald with nitpicky comments, and Tiffany demoralized him.*
8. *Miguel ruthlessly taunted Daniel, and Victoria harassed him.*
9. *Brennan played Kevin in pool, and Sara beat him at darts.*
10. *Lara approached Marina about the job, and Christina wrote her.*
11. *Amanda wrote Alisha, and Stevie made amends with her.*
12. *Amy met Kaitlin, and Hope saw her in person for the first time.*

13. *Kylie took Larissa to the waterpark, and Spencer went to a movie with her.*

14. *May picked Allison to be on the dodgeball team, and Gerald chose her as captain of the hockey team.*

15. *Dale sold Noah insurance, and Steven asked him to sign a petition.*

16. *Samuel threatened Jeff with a knife, and Erin Blindfolded him*

17. *Austin nudged Joseph, and Wesley bumped him.*

18. *Lorna phoned Molly, and Mariah emailed her.*

19. *Nelly stabbed Jacqueline, and Maya kicked her.*

20. *Sally called Monica, and Robert summoned her.*

21. *Dallas goaded Nathaniel, and Brianna ridiculed him.*

22. *Brittany elbowed Breanna, and Jason checked her.*

23. *Chase videotaped Marcus, just like Juliana photographed him.*

24. *Cole deceived Matthew, and Tasha tricked him.*

25. *Ida sewed Tammy a jacket, and Terry knit her a quilt.*

26. *Bryce made Liam a sandwich, and Thomas poured him a glass of juice.*

27. *Madison passed Shayla on a bike, and Savannah passed her in a car.*

28. *Jen wrote Margaret a song, and Ella wrote her a poem.*

29. *Alyssa failed Lisa on the chemistry lab, and Sophia failed her on the poetry assignment.*

B.2 SPANISH EXPERIMENT 1B, EQUIBIASED

1. *Alejandro parodió a Adolfo y Jaime lo castigó a él.*
2. *Alex desalentó a Amado con comentarios quisquillosos y Teresa lo desmoralizó a él.*
3. *Alondra reprobó a Patricia en laboratorio y Susana la reprobó en poesía a ella.*
4. *Amaranta conoció a Violeta en la escuela y Esperanza la vio en persona por primera vez a ella.*
5. *Belinda codeó a Brisa y Julio la golpeó a ella.*
6. *Bernardo le hizó un sandwich a Luis e Israel le sirvió un vaso con jugo a él.*
7. *Bersaín videograbó a Marcos y Josefina lo fotografeó a él.*
8. *Carlos ridiculizó a Jeffrey con crueldad y Emma lo imitó a él.*
9. *Celia eligió a Luz como parte de su equipo de futbol y Gilberto la escogió como capitán del equipo de voleibol a ella.*
10. *Dalinda llamó a Micaela y Rosa le escribió a ella.*
11. *Darío le vendío un seguro a Noé y Abel le pidió firmar una petición a él.*
12. *Diana llevó al parque a Ofelia y Simón la llevó al cine a ella.*
13. *Elda cosió un saco a Dalia y Guillermina le tejió un edredón a ella.*
14. *Evelia decepcionó a Tania y Matías la engañó a ella.*
15. *Héctor brindó por Ronaldo y Ana le dio una serenata a él.*
16. *Humberto molestó a Pepe con preguntas irrespetuosas y Quique lo irritó a él.*
17. *Javier atormentó a Manuel y Lisa lo fastidió a él.*

18. *Juan de Dios engañó a Nicolás y Cristal lo estafó a él.*
19. *Lara se dirigió a Mariana por el trabajo y Cristina le escribió a ella.*
20. *Matilda rebasó a Sofía con la bicicleta y Santiago la rebasó con el carro a ella.*
21. *Nelly acuchilló a Jaqueline e Inés la pateó a ella.*
22. *Rebeca le dio a Adalia una advertencia y Carlota le dijo que actuara con propiedad a ella.*
23. *Rocío le escribió una canción a Margarita y Pilar le escribió un poema a ella.*
24. *Samuel amenazó a Emanuel con un cuchillo y Erika le vendó los ojos a él.*
25. *Sandra llamó a Mónica y Rogelio la citó a ella.*
26. *Zacarías lastimó seriamente a Fernando y Carolina le hizo un moretón a él.*
27. *Agustín codeó a Jaime y William lo empujó a él.*

B.3 ENGLISH EXPERIMENT 2A

B.3.1 *Subject-biased*

1. *Alexander parodied John, and James chastized him.*
2. *Bob registered Justin, and Scott asked him for the registration fee.*
3. *Andrew fouled Jonah, and Calista ejected him from the game.*
4. *Jenna visited Ivy in hospital, and Tyson gave her a package to pass on.*
5. *Nicole yelled at Charlotte, and Isaiah pleaded with her to relax.*
6. *Alyssa annoyed Julie, and Katherine sat her out.*

7. *Colin taught Maxwell bad driving habits, and Ethan got angry at him.*

8. *Emily escaped from Justine, and Shane allowed her to stay in his attic.*

9. *Brandon cut Kenneth, and Carolina scolded him.*

10. *Michelle coached Brittney, and Christine paid her.*

11. *Olivia broke Tamara's arm, and Eva charged her with assault.*

12. *Kayla caught Lindsay stealing, and Heather thanked her.*

13. *Sabrina bullied Veronica, and Jane reported her to the teacher.*

14. *Sarah gave a cookie to Kate, and Victor appreciated her kindness*

15. *Dustin recruited Philip, and Tanner sent him the finder's fee.*

16. *Jennifer caught Alexa, and Grace helped her alert the authorities.*

B.3.2 *Object-biased*

1. *Lasha described Marika, and Joanna recognized her.*

2. *Rebecca pointed Carmen out, and Kimberley saw her.*

3. *Samantha misinformed Marissa, and Laura informed her of the truth.*

4. *Jessica burnt Molly, and Christie bandaged her.*

5. *Rachel wound Kristen up, and Sabrina quieted her down.*

6. *Becky gave Bethany a warning, and Courtney told her to act proper.*

7. *Giselle complained Naomi didn't work, and Vincent fired her.*

8. *Calvin upset Lucas, and Ryan calmed him down.*
9. *Dawson tricked Owen, and Mikayla had to tell him about the lie.*
10. *Lauren discovered Marina breaking a chair, and Nowlton disciplined her.*
11. *Sydney invited Vivian over for dinner, and Evan gave her the address.*
12. *Alexandra woke Molly up, and Zoe had to put her back to bed.*

B.4 SPANISH EXPERIMENT 2B

B.4.1 *Subject-biased*

1. *Amanda le escribió a Elisa y Esteban la compensó a ella.*
2. *Andrés fauleó a Rubén y Carolina lo expulsó del juego a él.*
3. *Beto provocó a Daniel y Adela lo ridiculizó a él.*
4. *Carlos le enseñó a Martín malos hábitos de manejo y Ernesto le gritó a él.*
5. *Cassandra le enseño inglés a Katia y Enrique le dio las gracias a ella.*
6. *Darío contrató a Pedro y Tomás le mandó la cuenta de honorarios a el.*
7. *David ayudó a Oscar a mudarse de casa y Leo pensó que era una persona servicial a él.*
8. *Eduardo acusó a Norberto de malversación y Verónica le exigió a él las razones por las cuales realizaba la acusación.*
9. *Jenny visitó a Irina en el hospital y Cuauhtémoc le dio un paquete para que se lo entregará a ella.*
10. *Karina atrapó a Hilda robando y Leonora le dio las gracias a ella.*

11. *Michelle entrenó a Bertha y Ana María le pagó a ella.*
12. *Miguel se burló despiadamente de Fabian y Victoria lo acusó a él.*
13. *Nicole le gritó a Beatriz y Lorena le pidió que se calmara a ella.*
14. *Olivia le rompió el brazo a Tamara y Eva la acusó de agresión a ella.*
15. *Sabrina molestó a Veronica y Juana la reportó con la maestra a ella.*
16. *Sara le dio una galleta a Karen y Victor le dijo que era muy amable a ella.*

B.4.2 *Object-biased*

1. *Alejandra despertó a Martha y Xiomara la tuvo que dormir de nuevo a ella.*
2. *Brandon cortó a Renato y Carolina lo vendó a él.*
3. *Calvin molestó a Lucas y Manolo lo calmó a él.*
4. *Darío engañó a Paco y Miguel le confesó la mentira a él.*
5. *Gisela se quejó de que Noemi no trabajaba y Vicente la corrió a ella.*
6. *Jessica quemó a Andrea y Leti la vendó a ella.*
7. *Josefa atrapó a Alba y Fátima la ayudó a escapar a ella.*
8. *Larisa describió a Marisela y Joana la reconoció a ella.*
9. *Laura descubrió a Marina rompiendo una silla y Tito la castigó a ella.*
10. *Natalia señaló a Carmén y Rosalva la vió a ella.*
11. *Raquel hirió a Crisitina y Sabine la acalló a ella.*
12. *Roberto registró a José y Gustavo le enseño a él.*

13. *Samanta mal informó a Maritza y Edith le informó la verdad a ella.*

14. *Sandra invitó a comer a Viridiana y Lydia le dio la dirección a ella.*

C

APPENDIX: CHAPTER 4 STIMULI

C.1 DESCRIPTION OF STIMULI

For clarity of presentation and reasons of space the plausible and implausible version of the introduction are presented together, the plausible version is always first in *one font* and the implausible always second, after the slash (/) in another font. Following this, we present the acknowledgment (*mhm* or *oh*), and finally the critical sentence. A: indicates one speaker and B: another.

C.2 STIMULI

1. A: *Ferdinand was een politicus die helemaal niet kon debatteren en Annelie kon hem goed nadoen.* / Ferdinand was een populaire politicus die erg overtuigend kon debatteren en Annelie was zijn grootste fan. B: *Mhm.* A: *Annelie bespotte hem.*

2. A: *Bas gooide met verfbommetjes toen de studenten aan het demonstreren waren.* / Bas stond op een afstandje toe te kijken toen de studenten aan het demonstreren waren. B: *Oh.* A: *De agent overmeesterde hem.*

3. A: *Anke had echt de grootste tas van alle feestgangers en moest voor de ingang van de discotheek wachten.* / Anke was als hoofdagent verantwoordelijk voor de feestgangers die voor de ingang van de discotheek wachten. B: *Mhm.* A: *De beveiliger controleerde haar.*

4. A: *André was een super getalenteerde jongleur in het kindercircus en wilde graag voor het grote publiek optreden.* / André was een bijzonder saaie jongleur in het kindercircus en wilde met zijn suffe act publiek trekken. B: *Oh.* A: *De circusdirecteur filmde hem.*

5. A: *Sofie deed steeds haar best bij het voorbereiden van de toets maar was de enige die in de les nooit vragen stelde.* / Sofie haalde het slechtste cijfer voor de toets en was altijd de enige die in de les geen enkele vraag stelde. B: *Mhm.* A: *De docent onderschatte haar.*

6. A: *Mieke is de beste leerling van de klas maar maakte een piepklein foutje in het dictee.* / Mieke is de beste leerling van de klas en had geen enkele fout in haar opstel gemaakt. B: *Oh.* A: *De meester verbeterde haar.*

7. A: *Bram droeg tijdens de voorlichtingsdag van de kunstacademie een fantasieloze outfit.* / Bram droeg tijdens de voorlichtingsdag van de kunstacademie een prachtig outfit. B: *Mhm.* A: *De ontwerper prees hem.*

8. A: *De degelijke Betty werkte als fiscaal agent aan de fraudezaak van de gepensioneerde miljonair Diederick.* / De degelijke Betty werkte als secretaresse in de zaak van de gepensioneerde miljonair Diederick. B: *Oh.* A: *Diederick huwde haar.*

9. A: *Toen Gerard's verloofde Christien aan de drank raakte, emigreerde hij en stichtte een nieuw gezin.* / Christien was Gerards meest geliefde dochter toen hij emigreerde en een nieuw gezin stichtte. B: *Mhm.* A: *Gerard verzweeg haar.*

10. A: *De agent was heel vervelend geweest en had Hilbert - die een slagroomtaart bij zich had - niet doorgelaten.* / De agent zag er keurig uit en was zo aardig geweest Hilbert - die oranjetompoucen bij zich had - gewoon door te laten. B: *Oh.* A: *Hilbert besmeurde hem.*

11. A: *Bastiaan was volstrekt geloofwaardig toen hij de goedgelovige Ida vertelde dat hij een boodschap van God had gekregen.* / Bastiaan was niet erg overtuigend toen hij de sceptische Ida vertelde dat hij een boodschap van God had gekregen. B: *Mhm.* A: *Ida vereerde hem.*

12. A: *Ruben was een vervelende werknemer van Ilse die aan alle vrouwen van het bedrijf schunnige brieven schreef.* / Ruben was een aardige werknemer van wie Ilse volgens haar verhalen goede rapporten kreeg. B: *Oh.* A: *Ilse ontsloeg hem.*

13. A: *Rens hoorde bij een groepje pestkoppen waar iedereen wel een beetje een hekel aan had en daar liep Iris recht op af.* / Rens hoorde bij een groepje stoere jongens waar iedereen wel een beetje verliefd op was en daar liep Iris recht op af. B: *Mhm.* A: *Iris omarmde hem.*

14. A: *Diane was de vechtpartij begonnen en het kon Jim niets schelen of er rake klappen vielen.* / Diane was de vechtpartij begonnen maar toch probeerde Jim te voorkomen dat er rake klappen vielen. B: *Oh.* A: *Jim beschermde haar.*

15. A: *Ricky was de enige leerling die Engels niet als moedertaal had. Juf Jelly daarentegen was in London geboren.* / Ricky was de enige leerling die Engels als moedertaal had. Juf Jelly daarentegen had pas een maand Engelse les. B: *Mhm.* A: *Jelly corrigeerde hem.*

16. A: *Die lieve Jelmer was met Judith's fiets door het glas gereden en bood aan om de band te plakken.* / Die nare Jelmer had met opzet Judith's fiets gemold en moest voor straf de band van haar fiets plakken. B: *Oh.* A: *Judith hielp hem.*

17. A: *Thera, het allerleukste meisje van de klas, was goed te zien toen Karin op de mensenmenigte toestapte.* / Thera, het onvriendelijkste meisje van de klas, stond heel ver weg, en Karin had geen zin om sociaal te doen. B: *Mhm.* A: *Karin groette haar.*

18. A: *Amy was wel de ergste van de meiden die Kevin tot het uiterste dreven.* / Amy was wel de leukste van de meiden die hevig op Kevin verliefd waren. B: *Oh.* A: *Kevin sloeg haar.*

19. A: *Mike was net zo dol op zijn kat als alle anderen en huilde bij Kim toen het beestje stierf.* / Mike was helemaal niet zo dol op Kims kat als alle anderen en negeerde Kim toen het beestje stierf. B: *Mhm.* A: *Kim condoleerde hem.*

20. A: *De oom van Kira houdt erg van flauwe 1-aprilgrappen en krijgt dan vaak de slappe lach.* / De oom van Kira heeft een soort zesde zintuig voor flauwe 1-aprilgrappen en is er nog nooit ingetrapt. B: *Oh.* A: *Kira fopte hem.*

21. A: *Daan was directeur van een kerncentrale en kwam in het winkelcentrum Leanne tegen die een topspion was van een buitenlandse inlichtingendienst.* / Daan was al jaren dakloos en kwam in het winkelcentrum Leanne tegen die bij de huisartsenpost werkte. B: *Mhm.* A: *Leanne bespioneerde hem.*

22. A: *Aan tafel kon Els maar niet stoppen met praten over de afgang van Maarten* / Aan tafel vertelde Els honderduit over het succes van Maarten. B: *Oh.* A: *Maarten schopte haar.*

23. A: *Frans had tijdens de workshop veel kritiek op Marieke waardoor ze onverwacht agressief werd.* / Frans gaf tijdens de workshop complimentjes aan Marieke waar ze zichtbaar van genoot. B: *Mhm.* A: *Marieke sloeg hem.*

24. A: *In de supermarkt was Thomas de enige collega met wie Melissa uren kon spreken over de klanten aan de kassa.* / In de supermarkt was Thomas de enige collega van Melissa die haar kon verbijsteren door uren te praten over quantum mechanica. B: *Oh.* A: *Melissa begreep hem.*

25. A: *Ahmed sprak Nederlands toen hij met de Nederlandse Merel zijn ouders in de Bijlmer bezocht.* / Ahmed sprak een onbekend dialect toen hij met de Nederlandse Merel zijn ouders in

Marokko bezocht. B: *Mhm.* A: *Merel verstond hem.*

26. A: *Lindsey had het slankste postuur van alle deelnemers en jurylid Mik moest beslissen wie Miss Holland werd.* / Lindsey had het lelijkste gezicht van alle deelnemers en jurylid Mik moest beslissen wie Miss Holland werd. B: *Oh.* A: *Mik verkoos haar.*

27. A: *Wanneer Neeltje haar drugs niet meer kon betalen, ging ze vaak even bij haar vader logeren* / Als Neeltje in de problemen zat, werd ze altijd door haar vader geholpen. B: *Mhm.* A: *Neeltje beroofde hem.*

28. A: *Rebecca was behoorlijk gewelddadig en daardoor besloot Nick om niet te gaan samenwonen.* / Rebecca was de liefde van zijn leven en daardoor besloot Nick om te gaan samenwonen. B: *Oh.* A: *Nick verliet haar.*

29. A: *Jansje woonde in een studenten flat toen ze een hartaanval kreeg en noodarts Piet wist dat jongeren vaak veel baat bij reanimatie hebben.* / Jansje woonde in een bejaardenflat toen ze een hartaanval kreeg en noodarts Piet wist dat ze niet bijgebracht wilde worden. B: *Mhm.* A: *Piet reanimeerde haar.*

30. A: *Sander had in de strijd een mes bemachtigd en liep dreigend naar Rianne die een jachtgeweer had.* / Sander had in de strijd een vuurwapen bemachtigd en verdedigde Rianne die alleen een broodmes had. B: *Oh.* A: *Rianne vermoordde hem.*

31. A: *Richards moeder was een uitzonderlijk liefdevolle opvoeder geweest.* / Richards moeder was geen liefdevolle opvoeder en ook geen goed voorbeeld voor haar zoon geweest. B: *Mhm.* A: *Richard vereerde haar.*

32. A: *Miranda draagt veel meer bij dan de andere secretaresses en Roland vindt haar een erg competente collega.* / Miranda draagt veel minder bij dan de andere secretaresses en Roland vindt haar een erg incompetente collega. B: *Oh.* A: *Roland loofde haar.*

33. A: *Marja had zich als enige van de kandidaten asociaal gedragen toen ze voor het TV programma uitging met Ruud.* / Marja had het als enige van de kandidaten heel gezellig gehad toen ze voor het TV programma uitging met Ruud. B: *Mhm.* A: *Ruud negeerde haar.*

34. A: *Boudewijn was in alle opzichten een slechte match voor Selina zeker vergeleken met haar vorige partners.* / Boudewijn was in alle opzichten een goede match voor Selina, zeker vergeleken met haar vorige partners. B: *Oh.* A: *Selina dumpte hem.*

35. A: *Ina blijft 's avonds erg vaak bij Sjors eten en Sjors dekte de tafel voor te weinig personen.* / Ina blijft 's avonds bijna nooit bij Sjors eten en was voor Sjors in het volkstuintje aan het werken. B: *Mhm.* A: *Sjors vergat haar.*

36. A: *Anneke was de beste spion van de AIVD en maakte misbruik van haar man Stijn totdat het fout ging.* / Anneke was de beste spion van de AIVD en ze had haar opleiding te danken aan haar man Stijn. B: *Oh.* A: *Stijn verried haar.*

37. A: *Kees is de grootste kwajongen van het dorp en jatte bij de bakker laatst een appelflap.* / Kees is de braafste jongen van het dorp en hij keek bij de bakker verlekkerd naar een appelflap. B: *Mhm.* A: *De bakker achtervolgde hem.*

38. A: *Boris was de enige verdachte die tijdens het verhoor veel onbekende details over het delict kon beschrijven.* / Boris was de enige verdachte die tijdens het verhoor geen details over het delict kon beschrijven. B: *Oh.* A: *De agent arresteerde hem.*

39. A: *Lucas was echt onzichtbaar voor de overige verkeersdeelnemers toen hij zonder achterlicht naar huis fietste.* / Lucas was goed zichtbaar voor de overige verkeersdeelnemers toen hij met voor en achterlicht naar huis fietste. B: *Mhm.* A: *De agent beboette hem.*

40. A: *Tim vroeg op de speelplaats aan Lieke of hij nu eindelijk de step mocht hebben.* / Tim zei op de speelplaats tegen Lieke dat zij nu

de step mocht hebben. B: *Oh.* A: *Lieke beet hem.*

41. A: *De piepjonge Fiona was veruit de mooiste deelneemster aan de modellenwedstrijd.* / De hoogbejaarde Fiona had kennissen die meededen aan de modellenwedstrijd. B: *Mhm.* A: *Het jurylid bekeek haar.*

42. A: *Pim rende tijdens de voetbalwedstrijd buitengewoon ongemotiveerd achter de bal aan.* / Pim rende tijdens de voetbalwedstrijd heel gemotiveerd achter de bal aan. B: *Oh.* A: *De coach bekritiseerde hem.*

43. A: *Karlijn werkte veel harder dan haar collega's om het project op tijd af te krijgen.* / Karlijn werkte veel minder hard dan haar collega's om het project op tijd af te krijgen. B: *Mhm.* A: *De werkgever beloonde haar.*

44. A: *Jordy was duidelijk de meest competente sollicitant en Stella had dat haarfijn door.* / Jordy was duidelijk de minst competente sollicitant en Stella had dat haarfijn door. B: *Oh.* A: *Stella benoemde hem.*

45. A: *Marlies was de slordigste kamerbewoonster en hield zich nooit aan het schoonmaakrooster dat Jaap had opgesteld.* / Marlies was de netste kamerbewoonster en hield zich aan het schoonmaakrooster dat Jaap had opgesteld. B: *Mhm.* A: *Jaap berispte haar.*

46. A: *Renske bleef altijd met de andere leerlingen kletsen, ook al wilde leraar Karsten met de les beginnen.* / Renske stopte net als de andere leerlingen altijd meteen met kletsen, als leraar Karsten met de les wilde beginnen. B: *Oh.* A: *Karsten berispte haar.*

47. A: *Julia was eerder door haar ex-man Erik beschreven als een uitermate onbetrouwbaar iemand.* / Julia was eerder door haar ex-man Erik beschreven als een uitermate betrouwbaar iemand. B: *Mhm.* A: *Erik beschuldigde haar.*

48. A: *Gerben was helemaal niet bang om ziek te worden terwijl zijn collega's bij de verkouden Tineke uit de buurt bleven.* / Gerben was erg bang om ziek te worden en bleef net als zijn collega's bij de verkouden Tineke uit de buurt. B: *Oh.* A: *Tineke besmette hem.*

49. A: *Lise komt erg vaak te laat op school en vanochtend kwam ze een kwartier na de bel de school binnen.* / Lise komt bijna nooit te laat op school en vanochtend kwam ze nog geen minuutje na de bel de school binnen. B: *Mhm.* A: *De conciërge bestrafte haar.*

50. A: *Pieter had jaren bij de plantsoenendienst gewerkt en wist niet hoe hij Maria voor de gek kon houden zonder gezien te worden.* / Pieter had jaren bij de recherche gewerkt en wist dus hoe hij Maria voor de gek kon houden zonder gezien te worden. B: *Oh.* A: *Maria betrapte hem.*

51. A: *De gevangen tijger liep grommend in de kooi heen en weer toen Robin het kamp van de stropers ontdekte.* / De gevaarlijke tijger liep grommend in de kooi heen en weer toen Robin de dierentuin bezocht. B: *Mhm.* A: *Robin bevrijdde haar.*

52. A: *Barry lag al een hele tijd in het ziekenhuis nadat hij Peter die met een mes werd bedreigd had gered.* / Barry zat al een hele tijd in de gevangenis omdat hij Peter onder bedreiging van een mes had beroofd. B: *Oh.* A: *Peter bezocht hem.*

53. A: *Rudy gebruikte voor zijn dissertatie alleen betrouwbare gegevens en Charlotte had hem daar al vaker voor geprezen.* / Rudy gebruikte voor zijn dissertatie zelfverzonnen gegevens en Charlotte had hem daar al vaker op betrapt. B: *Mhm.* A: *Charlotte citeerde hem.*

54. A: *Nelly werd heel snel moe en kon de wandelgroep van haar vriend Kees niet meer bijhouden.* / Nelly werd helemaal niet snel moe en kon de wandelgroep van haar vriend Kees gemakkelijk bijhouden. B: *Oh.* A: *Kees droeg haar.*

55. A: *Maartje was extreem ongeduldig toen ze voor de achtbaan in de rij stond en probeerde zich voor Christa te dringen.* / Maartje was helemaal niet ongeduldig toen ze voor de achtbaan in de rij stond en wachtte met Christa op haar beurt. B: *Mhm.* A: *Christa duwde haar.*

56. A: *Anita zag dat haar man Mark na het feest meteen naar bed ging en zich niet om de troep bekommerde.* / Anita zag dat haar man Mark na het feest de troep aan het opruimen was. B: *Oh.* A: *Mark ergerde haar.*

57. A: *Nadia kreeg net als haar buren de postcodeprijs uit handen van Jolanda.* / Nadia kreeg in tegenstelling tot haar buren geen postcodeprijs uitgereikt door Jolanda. B: *Mhm.* A: *Jolanda feliciteerde haar.*

58. A: *Anne speelde vreselijk goed en won de finale van de schaakwedstrijd.* / Anne speelde vreselijk slecht en verloor de finale van de schaakwedstrijd. B: *Oh.* A: *De juf feliciteerde haar.*

59. A: *Van zijn vrienden was Freek de enige met zo'n stoer alternatief kapsel dat voorbijgangers hem altijd nakeken.* / Van zijn vrienden was Freek de enige met zo'n saai kapsel dat voorbijgangers hem nooit opmerkten. B: *Mhm.* A: *De trendscout fotografeerde hem.*

60. A: *Liesbeth wist dat Jonathan - net als de rest van zijn familie - verschrikkelijk eerlijk was. Jonathan zei dat hij het niet had gedaan.* / Liesbeth wist dat Jonathan - net als de rest van zijn familie - een verschrikkelijk liegbeest was. Jonathan zei dat hij het niet had gedaan. B: *Oh.* A: *Liesbeth geloofde hem.*

61. A: *Chantal was 's werelds gemeenste ex-vriendin en verspreidde allerlei roddels over Sander.* / Chantal was 's werelds liefste ex-vriendin en ontkende alle roddels over Sander. B: *Mhm.* A: *Sander haatte haar.*

62. A: *Henk leek nog volkomen op zijn vroegere zelf en zijn schooljuf Hennie liep hem stralend tegemoet op de reunie.* / Henk leek in niets op zijn vroegere zelf en zijn schooljuf Hennie liep hem straal

voorbij op de reunie. B: *Oh.* A: *Hennie herkende hem.*

63. A: *Lisette hield zich nooit in met gitaar spelen omdat ze dacht dat buurman Ramon zo goed als doof was.* / Lisette hield zich altijd in met gitaar spelen ook al wist ze dat buurman Ramon zo goed als doof was. B: *Mhm.* A: *Ramon hoorde haar.*

64. A: *Dennis maakte sinds zijn relatie met Lindsey creatievere schilderijen dan zijn collega's.* / Dennis maakte sinds zijn relatie met Lindsey eentonigere schilderijen dan zijn collega's. B: *Oh.* A: *Lindsey inspireerde hem.*

65. A: *Nienke was de hele ochtend vervelend geweest op de crêche, en had haar koekje niet met Nico gedeeld.* / Nienke was de hele ochtend braaf geweest op de crêche, en had haar koekje met Nico gedeeld. B: *Mhm.* A: *Nico kneep haar.*

66. A: *Sandra was erg sportief toen Tom tijdens de gymles spelers voor het korfbalteam zocht.* / Sandra kon geen bal vangen toen Tom tijdens de gymles spelers voor het korfbalteam zocht. B: *Oh.* A: *Tom koos haar.*

67. A: *Chris tekende van alle straatartiesten het best gelijkende portret van Elly.* / Chris tekende van alle straatartiesten het minst gelijkende portret van Elly. B: *Mhm.* A: *Elly loofde hem.*

68. A: *Berry ging met wat studiegenoten in het buitenland studeren en was volgens Lynn erg knap en charmant.* / Berry ging met wat studiegenoten in het buitenland studeren en was volgens Lynn erg vervelend en opdringerig. B: *Oh.* A: *Lynn miste hem.*

69. A: *Jos bedacht de meest geniale uitvinding en vroeg Sanne om sponsorgeld voor de uitvoering ervan.* / Jos bedacht de meest nutteloze uitvinding en vroeg Sanne om sponsorgeld voor de uitvoering ervan. B: *Mhm.* A: *Sanne ondersteunde hem.*

70. A: *Katja was overduidelijk betrokken bij de zaak waar Koen onderzoek naar deed.* / Katja was overduidelijk niet betrokken bij

de zaak waar Koen onderzoek naar deed. B: *Oh.* A: *Koen ondervroeg haar.*

71. A: *Klaas brak zijn been toen hij met zijn racefiets tegen een geparkeerde vrachtwagen reed.* / Klaas had geen schram toen hij met zijn vrachtwagen tegen een geparkeerde fiets reed. B: *Mhm.* A: *De arts onderzocht hem.*

72. A: *Britt heeft als enige van de kinderen een homohuwelijk gesloten, ook al was vader Jan het daar totaal niet mee eens.* / Britt heeft als enige van de kinderen een gewoon huwelijk gesloten, ook al had vader Jan dat helemaal niet meer verwacht. B: *Oh.* A: *Jan ontierf haar.*

73. A: *Toen de enigszins labiele Roland vertelde dat hij van alle sollicitanten het hoogste was geeïndigd, was Ingrid zeer bezorgd.* / Toen een stralende Roland vertelde dat hij van alle sollicitanten het hoogste was geeïndigd, was Ingrid zeer verheugd. B: *Mhm.* A: *Ingrid ontmoedigde hem.*

74. A: *Marianne was erg emotioneel toen Rogier in het overvolle restaurant voor haar neerknielde.* / Marianne was erg boos toen Rogier in het overvolle restaurant haar plaats niet bezet wilde houden. B: *Oh.* A: *Rogier ontroerde haar.*

75. A: *Martine werkt van al haar collega's het minst efficient, dus haar baas Bart was zeer ontevreden.* / Martine werkt van al haar collega's het meest efficient, dus haar baas Bart was dik tevreden. B: *Mhm.* A: *Bart ontsloeg haar.*

76. A: *Frederik was ondanks de crisis de beste zakenman van het jaar, dus crimineel Klaas verwachtte veel losgeld voor hem te kunnen eisen.* / Frederik was net als veel ondernemers door de crisis failliet gegaan, dus crimineel Klaas verwachtte geen geld meer voor hem te kunnen eisen. B: *Oh.* A: *Klaas ontvoerde hem.*

77. A: *Uit de onderzoeken van plastisch chirurg Guido was gebleken dat met name Stefanie nog een operatie nodig had.* / Uit de onderzoeken van plastisch chirurg Guido was gebleken dat met name

Stefanie nog een operatie niet zou overleven. B: *Mhm.* A: *Guido opereerde haar.*

78. A: *Sabine was niet zo heel erg slim, haalde de meeste onvoldoendes, maar werd toch voorwaardelijk tot de studie toegelaten.* / Sabine was niet zo heel erg slim, haalde de meeste onvoldoendes, en werd uiteindelijk toch niet tot de studie toegelaten. B: *Oh.* A: *De studieadviseur overschatte haar.*

79. A: *Juf Inge heeft volgens mij de zachtste stem van iedereen die ik ken, en ze is heel verlegen.* / Juf Inge heeft volgens mij de hardste stem van iedereen die ik ken, en ze praat voortdurend. B: *Mhm.* A: *De voorleesmoeder overstemde haar.*

80. A: *Arnold was de kleinste van de klas, terwijl Janine juist groot voor haar leeftijd is.* / Arnold was de sterkste van de klas, terwijl Janine juist een erg tenger meisje is. B: *Oh.* A: *Janine pestte hem.*

81. A: *Willemijn was een verlegen meisje dat was blijven zitten en iedereen keek neer op haar, ook Jord.* / Willemijn was een wilde meid die was blijven zitten en iedereen was bang voor haar, ook Jord. B: *Mhm.* A: *Jord plaagde haar.*

82. A: *Wendy is erg goed in pottenbakken en haar man David was erg onder de indruk, zo zei hij tegen zijn vrienden.* / Wendy is erg slecht in pottenbakken en haar man David schaamde zich dood, zo zei hij tegen zijn vrienden. B: *Oh.* A: *David prees haar.*

83. A: *Patrick was de beroerdste zwemmer van al zijn vrienden en strandwacht Lara zag hoe hij gillend en schreeuwend in zee lag.* / Patrick was de beste zwemmer van al zijn vrienden en strandwacht Lara zag hoe hij soepel en gespierd de zee in ging. B: *Mhm.* A: *Lara redde hem.*

84. A: *Kapitein Rik van de verzetsgroep Noord redde veel levens door zijn moedige gedrag tijdens de oorlog.* / Kapitein Rik van de verzetsgroep Noord bracht veel levens in gevaar door zijn overmoedige gedrag tijdens de oorlog. B: *Oh.* A: *De Koningin rid-*

derde hem.

85. A: *Jolien had deze keer maar weer eens parfum opgedaan, hoewel ze wist dat Emiel daar een hekel aan had.* / Jolien had deze keer maar eens geen parfum opgedaan, omdat ze wist dat Emiel daar een hekel aan had. B: *Mhm.* A: *Emiel rook haar.*

86. A: *Leon kon heel goed stilzitten en was het perfecte model voor Rosalie en haar schildervrienden.* / Leon kon helemaal niet stilzitten en was dus geen optie om model te zijn voor Rosalie en haar schildervrienden. B: *Oh.* A: *Rosalie schilderde hem.*

87. A: *Lieneke was volgens Simon heel erg irritant toen ze tijdens het familiediner niet ophield over opa te praten.* / Lieneke was volgens Simon heel erg begrijpend toen hij tijdens het familiediner niet ophield over opa te praten. B: *Mhm.* A: *Simon schopte haar.*

88. A: *Max en zijn familie waren allergisch voor katten en Sylvia wou de garage daarom niet als asiel, maar als schoonheidssalon inrichten.* / Max en zijn familie waren allergisch voor katten en Sylvia wou de garage desondanks niet als schoonheidssalon, maar als asiel inrichten. B: *Oh.* A: *Max steunde haar.*

89. A: *Hans kan heel blij worden als je hem uit zijn dagelijkse werkzaamheden haalt, en dat is iets wat Eline heel goed begrijpt.* / Hans kan heel boos worden als je hem uit zijn dagelijkse werkzaamheden haalt, en dat is iets wat Eline heel goed begrijpt. B: *Mhm.* A: *Eline stoorde hem.*

90. A: *Lisa had alle foto's geruïneerd om het Vincent en zijn collega's moeilijk te maken.* / Lisa had alle foto's gerubriceerd om het Vincent en zijn collega's gemakkelijk te maken. B: *Oh.* A: *Vincent strafte haar.*

91. A: *Geert is de allerbeste monteur en zijn baas Theo kreeg een belangrijke klus van een miljoenenbedrijf.* / Geert is de slechtste monteur en zijn baas Theo kreeg een belangrijke klus van een miljoenenbedrijf. B: *Mhm.* A: *Theo stuurde hem.*

92. A: *Harm begrijpt instructies veel sneller dan andere studenten en Elise had alleen slimme proefpersonen nodig.* / Harm begrijpt instructies veel trager dan andere studenten en Elise had alleen slimme proefpersonen nodig. B: *Oh.* A: *Elise testte hem.*

93. A: *Nina kon niet echt snel rennen en was bij tikkertje zelfs de klos bij Sam die op krukken loopt.* / Nina kon heel snel rennen en was bij tikkertje geen partij voor Sam die op krukken loopt. B: *Mhm.* A: *Sam tikte haar.*

94. A: *Karels tante werd net als de andere familieleden erg emotioneel tijdens de begrafenis.* / Karels tante toonde zich als enige van de familieleden zelfs enigszins vrolijk tijdens de begrafenis. B: *Oh.* A: *Karel troostte haar.*

95. A: *Malou zorgde, in tegenstelling tot haar broers, helemaal niet voor haar bejaarde ouders en Michiel vond dat belachelijk.* / Malou zorgde, net als haar broers, elke dag voor haar bejaarde ouders en Michiel vond dat fantastisch. B: *Mhm.* A: *Michiel verachtte haar.*

96. A: *Na het feestje bij Daniël - eigenlijk geen echte vriend van Leonie - was ze haar merkhorloge kwijt.* / Na het feestje bij Daniël - de oprechtste vriend van Leonie - was ze haar merkhorloge kwijt. B: *Oh.* A: *Leonie verdacht hem.*

97. A: *Jenny kwam altijd het meest voor Florian op en de hele klas moest met de directeur over het pesten praten.* / Jenny pestte Florian altijd heel erg en de hele klas moest met de directeur over het pesten praten. B: *Mhm.* A: *Florian verdedigde haar.*

98. A: *Karina was vreselijk pijngevoelig en vroeg of de tandarts haar zere kies weg kon halen.* / Karina was altijd heel stoer bij de tandarts en vroeg of hij snel wat tandsteen weg kon halen. B: *Oh.* A: *De tandarts verdoofde haar.*

99. A: *Koen was een bijzonder gespierde sportschoolfanaat en Cindy vond sterke jongens aantrekkelijk.* / Koen was een magere jongen met een bril en Cindy vond juist sterke en stoere jongens

aantrekkelijk. B: *Mhm.* A: *Cindy verleidde hem.*

100. A: *Arjan was de klunzigste monteur van de machinefabriek en Arie, zijn chef, was een ongevoelige tiran.* / Arjan was de meest ervaren monteur van de machinefabriek en Arie, zijn chef, was een meevoelende coach. B: *Oh.* A: *Arie vernederde hem.*

101. A: *Marjan verwachtte helemaal geen bezoek, maar Luuk had iedereen uitgenodigd voor het feest.* / Marjan verwachtte een huis vol bezoek, en Luuk had inderdaad iedereen uitgenodigd voor het feest. B: *Mhm.* A: *Luuk verraste haar.*

102. A: *Willianne bezat alle kwaliteiten die Paul graag in vrouwen ziet.* / Willianne bezat geen van de kwaliteiten die Paul graag in vrouwen ziet. B: *Oh.* A: *Paul versierde haar.*

103. A: *Johan gooide de laagste score van de mannen en gaf toen de beurt aan Sylvia die tot dan elke week wel had gedart.* / Johan gooide de hoogste score van de mannen en gaf toen de beurt aan Sylvia die nog nooit had gedart. B: *Mhm.* A: *Sylvia versloeg hem.*

104. A: *Magische Marco was de slechtste tovenaar aller tijden en vluchtte voor zijn vijanden bij heks Johanna.* / Magische Marco was de machtigste tovenaar aller tijden en zocht naar zijn vijanden bij heks Johanna. B: *Oh.* A: *Johanna verstopte hem.*

105. A: *Christiaan bezwoer dat hij het nooit meer zou doen, en Carola was nog nooit door Christiaan bedrogen.* / Christiaan bezwoer dat hij het nooit meer zou doen, maar Carola was al drie keer eerder door Christiaan bedrogen. B: *Mhm.* A: *Carola vertrouwde hem.*

106. A: *Gisteren had Harry met zijn vrienden veel lol gemaakt op het feestje bij Marloes.* / Gisteren had Harry met zijn vrienden veel kapot gemaakt op het feestje bij Marloes. B: *Oh.* A: *Marloes verwelkomde hem.*

107. A: *Annie was de favoriete kleindochter van opa Gerrit.* / Annie was de minst favoriete zus van opa Gerrit. B: *Mhm.* A: *Gerrit verwende haar.*

108. A: *Anton was chronisch ziek en Antje besloot om haar baan op te zeggen en voortaan thuis te blijven.* / Anton kwam van vakantie terug en zag dat Antje en de kinderen zwaar verkouden waren. B: *Oh.* A: *Antje verzorgde hem.*

109. A: *Peuter Jeroen had vergeleken met de anderen bijna niets gegeten toen hun moeder Janneke met een voedzame stamppot binnenkwam.* / Peuter Jeroen had de meeste pannenkoeken gegeten van iedereen toen hun moeder Janneke met een voedzame stamppot binnenkwam. B: *Mhm.* A: *Janneke voedde hem.*

110. A: *Op het juniorenzeilkamp wilde Kirsten als enige van haar teamgenootjes tijdens het noodweer zeilen.* / Op het juniorenzeilkamp wilde Kirsten als enige van haar teamgenootjes niet tijdens het noodweer zeilen. B: *Oh.* A: *De coach waarschuwde haar.*

111. A: *Nicole wilde alles van religie weten en had een overduidelijke voorkeur voor godsdienstles.* / Nicole moest niets van religie weten en had een overduidelijke hekel aan godsdienstles. B: *Mhm.* A: *De priester zegende haar.*

112. A: *Anders dan haar oudere zussen was Ymke die avond niet op tijd thuis en hun vader Wouter was bezorgd.* / Net als haar oudere zussen was Ymke die avond op tijd thuis en hun vader Wouter was gerust. B: *Oh.* A: *Wouter zocht haar.*

113. A: *Dana haalde iedereen verkeerd in, en reed in een wak bij de wedstrijd op natuurijs.* / Dana haalde iedereen in, en reed een persoonlijk record op natuurijs. B: *Mhm.* A: *De trainer bekritiseerde haar.*

114. A: *Bertus bleek voor de subsidie in aanmerking te komen, en kwam bij Irene om het te vieren* / Bertus bleek niet voor de subsidie in aanmerking te komen, en kwam bij Irene om te mopperen B:

Oh. A: *Irene feliciteerde hem.*

115. A: *Heike was de enige die geen hulp aanbood toen Dirk in de problemen kwam.* / Heike was de enige die zijn hulp aanbood toen Dirk in de problemen kwam. B: *Mhm.* A: *Dirk haatte haar.*

116. A: *Mick doet veel meer dan de andere vakkenvullers en Bert vindt hem een waardevolle collega.* / Mick doet veel minder dan de andere vakkenvullers en Bert vindt hem een waardeloze collega. B: *Oh.* A: *Bert loofde hem.*

117. A: *Gijsbert had de hele week de minste omzet gehaald en Anja, zijn groepsleider, was daar niet blij mee.* / Gijsbert had de hele week de meeste omzet gehaald en Anja, zijn groepsleider, was daar erg blij mee. B: *Mhm.* A: *Anja ontsloeg hem.*

118. A: *Ronald hield ervan om in opvallende kleren rond te lopen, met grote kragen en wijde pijpen, net als Henrike.* / Ronald hield ervan om in opvallende kleren rond te lopen, met grote kragen en wijde pijpen, zeer tegen de zin van Henrike. B: *Oh.* A: *Henrike prees hem.*

119. A: *Betsie was de slordigste typiste van het hele kantoor en Jakob, de directeur, was bijzonder boos.* / Betsie was de beste typiste van het hele kantoor en Jakob, de directeur, was bijzonder trots. B: *Mhm.* A: *Jakob vernederde haar.*

120. A: *Adam was de enige broer van Bregje met wie ze nog steeds een goede band had.* / Adam was de enige broer van Bregje met wie ze nog steeds ruzie had. B: *Oh.* A: *Bregje verwende hem.*

SUMMARY

This thesis looked at anaphora resolution, concentrating in particular on the influence of markedness of anaphoric expressions (in terms of prosodic and/or morphological realization) and plausibility of the antecedents (based on world knowledge). One of the key issues in this type of investigation is how multiple sources of information are combined by the language processor. This was addressed by using a diverse set of psycholinguistic response measures (self-paced reading, antecedent selection and ERP (Event-Related brain Potentials) registration).

In Chapter 2, the results from the self-paced reading experiments suggested that in the comprehension of pronouns all sources of information are used in parallel, since the constraint-based model best described processing in the cases where participants made an effort to interpret the stories as soon as possible. Specifically, there was evidence of ambiguous items taking longer than unambiguous ones at the pronoun in the Antecedent Selection and the Plausibility Rating condition. There was no difference between ambiguous and unambiguous pronouns in the task condition that required no disambiguation. These results combined suggest that the constraint-based model should be modified in such a way that it can incorporate some kind of "task-engagement" parameter.

In Chapter 3, we looked at what information is used in off-line referent choice. There was evidence of a complex interaction between world knowledge, markedness, order of mention and syntax when selecting an antecedent, in an antecedent choice task. Markedness and world knowledge were shown to outweigh syntactic and order-of-mention cues. World knowledge was capable of dictating the referent to a strong degree. The results were different when pronouns were equiplausible, however. When there was no (world knowledge) bias towards either of two possible referents, marked delivery of the pronoun, either by accentuation or by adding morphological complexity, was able to switch the referent. When there was a strong bias toward one

antecedent, however, a marked pronoun reduced the preference for that antecedent, but did not switch it.

Chapter 4 examined the neural processes underlying the interaction of markedness and world knowledge in pronoun comprehension. Sentences where the only accessible referent for the pronoun was implausible and unaccented produced an early P600 effect, and not N400 associated with difficulty in comprehension. Accenting the pronoun produced an effect that could be interpreted as an Nref associated with increased effort applied to finding a plausible referent. Alternatively, it is also possible to interpret this effect as a reduced P600, perhaps indicating that interpretation did become easier by having an accent on the pronoun.

This research provides psycholinguistic and neurolinguistic evidence for the role of world knowledge and accentuation in pronoun processing, while also contributing to the ongoing discussion of the interpretation of ERP effects like the P600 and Nref.

SAMENVATTING

Dit proefschrift rapporteert onderzoek naar het verwerken van anaforen, en concentreert zich met name op de invloed van gemarkeerdheid van de anaforische uitdrukking (in termen van prosodische en/of morfologische kenmerken) en plausibiliteit (gebaseerd op kennis van de wereld). Een van de belangrijkste vragen in dit soort onderzoek is hoe de menselijke taalverwerker verschillende bronnen van informatie combineert bij het begrijpen van taal. Dit is onderzocht met behulp van uiteenlopende psycholinguistische methoden en paradigma's, zoals leestijdregistratie, off-line keuze taken, en ERP registratie (Event Related brain Potentials).

De resultaten van de leestijdexperimenten (zie Hoofdstuk 2) toonden aan dat bij het begrijpen van persoonlijke voornaamwoorden alle informatiebronnen in parallel gebruikt worden. In ieder geval was het zo dat er effecten van competitie waren (ambigue voornaamwoorden nemen meer tijd dan niet ambigue) wanneer de proefpersonen tot taak hadden de referent voor het voornaamwoord te rapporteren, of wanneer een plausibiliteitsoordeel werd gevraagd; er was geen evidentie voor een competitieeffect in de taakconditie waarin proefpersonen alleen maar hoefden te lezen. Samengenomen laten deze resultaten zien dat het Constraint-based Model, dat parallelle verwerking voorstaat, aangepast moet worden om ook aspecten van de specifieke taak die de lezer heeft te modelleren.

In Hoofdstuk 3 heb ik gekeken naar welke informatie gebruikt wordt bij de (off-line) keuze tussen twee mogelijke referenten van een ambigu voornaamwoord. Er was sprake van een complexe interactie tussen wereldkennis, gemarkeerdheid, 'order-of-mention' en syntaxis bij het kiezen van een referent. Gemarkeerdheid en wereldkennis wogen duidelijk zwaarder dan 'order-of-mention' of syntactische informatie. Wereldkennis was een uitermate sterke bepalende factor bij de keuze voor de ene of de andere referent. In de conditie waar er op basis van wereldkennis geen keuze gemaakt kon worden (d.w.z., waar beide ref-

erenten even plausibel waren) werd de keuze voor de referent gemaakt op basis van de gemarkeerdheid van de anaforische uitdrukking. Zowel het accentueren, als het toevoegen van morfologische complexiteit had tot gevolg dat de keuze voor de referent volledig omdraaide. Wanneer een van de referenten wel duidelijk meer plausibel was, werd de voorkeur door het markeren van het voornaamwoord wel aanzienlijk minder sterk, maar de voorkeur draaide niet om.

Hoofdstuk 4 beschrijft het onderzoek naar de neurale processen die aan het begrijpen van voornaamwoorden ten grondslag liggen. In het experiment werd met name gekeken naar de interactie tussen gemarkeerdheid en wereldkennis bij het begrijpen van voornaamwoorden. Implausibele zinnen waar het voornaamwoord niet geaccentueerd was lieten een vroege 'P600' zien (vergeleken met plausibele controle zinnen). Een P600 is een positieve verschuiving in het ERP signaal die vroeger geinterpreteerd werd als een teken van syntactische verwerkingsprocessen. De laatste tijd wordt de P600 steeds vaker gezien als een indicatie van problemen met het vormen van de betekenis van een zin. Er werd geen 'N400' (negatieve verschuiving in het ERP signaal) gevonden. Als het voornaamwoord wel geaccentueerd was, werd er een zogeheten Nref effect gevonden, een negativiteit die met name duidelijk aanwezig is op frontale electrodes. Deze Nref wordt doorgaans verbonden aan de moeite die het kost om bij een voornaamwoord de juiste referent te vinden. Het is echter ook mogelijk dit effect niet als een verhoogde negativiteit (Nref) te zien, maar als een verlaagde positiviteit, een gereduceerde P600. Misschien is de interpretatiemoeilijkheid verminderd doordat het voornaamwoord met de vreemde referent een accent heeft gekregen.

Dit onderzoek levert psycholinguïstische en neurolinguïstische evidentie voor de rol van wereldkennis en accentuering bij de verwerking van voornaamwoorden en draagt daarnaast bij aan de lopende discussie over de interpretatie van ERP effecten als de P600 en Nref.

RESUMEN

Esta tesis examina la resolución de la anáfora, concentrándose particularmente en la influencia de la forma marcada de las expresiones anafóricas (en términos de la forma prosódica y/o morfológica) y la verosimilitud de los antecedentes (basada en el conocimiento del mundo). Uno de los aspectos fundamentales de este tipo de investigación es cómo las múltiples fuentes de información se combinan en el procesamiento del lenguaje. Esto se investigó utilizando un amplio conjunto de instrumentos de medidas psicolingüísticas de respuesta (lectura autodirigida, selección de antecendentes y registro de ERP (Potenciales de Eventos Relacionados)).

En el capítulo 2, los resultados de los experimentos de lectura autodirigida sugieren que en la compresión de pronombres, todas las fuentes de información se utilizan en paralelo, ya que el modelo basado en restricción (el modelo Constraint-Based) describe mejor el procesamiento en los casos en los que los participantes hacen un esfuerzo para interpretar las historias tan pronto como sea posible. Específicamente, hubo evidencia que en la Selección de Antecedentes y el Grado de Clasificación de Verosimilitud en los pronombres los ítems ambiguos tomaron más tiempo que los ítems no ambiguos. En las tareas de condición que no requerían desambiguación no hubo diferencia entre pronombres ambiguos y no ambiguos. Estos resultados combinados sugieren que el modelo basado en restricción debe ser modificado de una manera que permita incorporar algún tipo de parámetro de "integración de tareas".

En el capítulo 3, observamos qué información es utilizada en la selección de referentes fuera de línea. Hubo evidencia de una interacción compleja entre conocimiento del mundo, marcación, orden de mención y sintaxis al momento de seleccionar un antecedente, en un instrumento de selección del antecedente. Se demostró que la marcación y el conocimiento del mundo tuvieron más influencia que la entrada sintáctica y el orden de mención. El conocimiento del mundo determinó en gran me-

dida el referente. Sin embargo, los resultados fueron diferentes cuando los referentes eran igualmente verosímiles. Cuando no había preferencia (del conocimiento del mundo) hacia ninguno de los dos posibles referentes, la marcación deliberada del pronombre, (ya sea por acentuación o por la complejidad morfológica), fue suficiente para cambiar el referente. Sin embargo, cuando había una preferencia, un pronombre marcado redujo la preferencia por ese antecedente, pero no lo cambió.

En el capítulo 4, examinamos el proceso neurológico subyacente a la interacción de la marcación y el conocimiento del mundo para la comprensión de pronombres. Los enunciados donde el único referente accesible para el pronombre era inverosímil y no acentuado crearon un efecto P600 temprano, y no un efecto N400 asociado con dificultad de comprensión. Acentuar el pronombre originó un efecto que podría ser interpretado como un Nref asociado a un incremento en el esfuerzo aplicado a encontrar un referente verosímil. De manera alternativa, también es posible interpretar este efecto como un reducido P600, indicando tal vez que la interpretación efectivamente se volvió más fácil al tener un acento en el pronombre.

Esta investigación provee evidencia psicolingüística y neurolingüística del rol que el conocimiento del mundo y la acentuación tiene en la comprensión de los pronombres, y contribuye también a la discusión actual de la interpretación de efectos ERP como el P600 y el Nref.

REFERENCES

Arnold, J. E. (1998). *Reference Form and Discourse Patterns*. Unpublished doctoral dissertation, Stanford University. (Cited on page 18.)

Arnold, J. E. (2010). How speakers refer: the role of accessibility. *Language and Linguistics Compass*, 4(4):187–203. (Cited on pages 18 and 40.)

Arnold, J. E. and Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, 56(4):521–536. (Cited on page 75.)

Arnold, J. E., Losongco, A., Wasow, T., and Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76(1):28–55. (Cited on page 43.)

Austen, J. (1994/n.d.). Pride and Prejudice. (Cited on page 51.)

Baauw, S., Ruigendijk, E., Cuetos, F., and Avrutin, S. (2011). The interpretation of stressed and non-stressed pronouns in spanish language breakdown. *Aphasiology*, 25(3):386–408. (Cited on pages 42 and 44.)

Barton, S. B. and Sanford, A. J. (1993). A case-study of anomaly detection: shallow semantic processing and cohesion establishment. *Memory and Cognition*, 21(4):477–487. (Cited on page 17.)

Blackwell, S. E. (2001). Testing the neo-Gricean pragmatic theory of anaphora: The influence of consistency constraints on interpretations of coreference in spanish. *Journal of Pragmatics*, 33(6):901–941. (Cited on pages 40, 42, 44, and 46.)

Boersma, P. and Weenink, D. (2009). Praat: doing phonetics by computer [computer program] (version 5.3.04). (Cited on page 84.)

Branigan, H. P., Pickering, M. J., Liversedge, S. P., Stewart, A. J., and Urbach, T. P. (1995). Syntactic priming: investigating the mental representation of language. *Journal of Psycholinguistic Research*, 24(6):489–506. (Cited on page 44.)

Breen, M., Fedorenko, E., Wagner, M., and Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, 25(7-9):1044–1098. (Cited on page 4.)

Brouwer, H., Fitz, H., and Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the p600 in language comprehension. *Brain Research*, 1446:127–43. (Cited on pages 9, 77, 78, 79, 80, 100, and 102.)

Brown, C. and Hagoort, P. (1993). The processing nature of the n400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5(1):34–44. (Cited on page 77.)

Burkhardt, P. (2007). The p600 reflects cost of new information in discourse memory. *NeuroReport*, 18(17):1851–1854. (Cited on pages 78, 79, 80, and 100.)

Callahan, S. M., Shapiro, L. P., and Love, T. (2010). Parallelism effects and verb activation: The sustained reactivation hypothesis. *Journal of Psycholinguistic Research*, 39(2):101–118. (Cited on page 44.)

Carrillo-de-la-Peña, M., Holguín, S. R., Corral, M., and Cadaveira, F. (1999). The effects of stimulus intensity and age on visual-evoked potentials (VEPs) in normal children. *Psychophysiology*, 36(6):693–698. (Cited on pages 80 and 104.)

Chafe, W. L. (1974). Language and consciousness. *Language*, 50(1):111–133. (Cited on page 71.)

Chambers, C. G. and Smyth, R. (1998). Structural parallelism and discourse coherence: A test of centering theory. *Journal of Memory and Language*, 39(4):593–608. (Cited on pages 3, 44, and 46.)

Christianson, K., Hollingworth, A., Halliwell, J., and Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42(4):368–407. (Cited on page 16.)

Crawley, R. A., Stevenson, R. J., and Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, 19(4):245–264. (Cited on pages 3, 40, 43, and 46.)

Daneman, M. and Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour*, 19(4):450–466. (Cited on page 18.)

De-la-Mota, C., Martín Butragueño, P., and Prieto, P. (2010). Mexican Spanish intonation. In *Transcription of Intonation of the Spanish language*, pages 319–350. Lincom Europa, Munich, Germany. (Cited on page 57.)

Dimitrova, D. V. (2012). *Neural Correlates of Prosody and Information Structure*. Doctoral dissertation, University of Groningen. (Cited on pages 8 and 80.)

Dimitrova, D. V., Redeker, G., and Hoeks, J. (2009). Did you say a BLUE banana? The prosody of contrast and abnormality in Bulgarian and Dutch. In *Tenth Annual Conference of the International Speech Communication Association*. (Cited on pages 4, 9, 71, 74, 99, and 110.)

Dimitrova, D. V., Stowe, L. A., Redeker, G., and Hoeks, J. C. (2012). Less is not more: neural responses to missing and superfluous accents in context. *Journal of Cognitive Neuroscience*, 24(12):2400–2418. (Cited on pages 80, 81, 91, and 104.)

Farmer, T. A., Cargill, S. A., Hindy, N. C., Dale, R., and Spivey, M. J. (2010). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science*, 31(5):889–909. (Cited on page 109.)

Ferreira, F., Bailey, K. G. D., and Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11(1):11–15. (Cited on page 16.)

Ferreira, F. and Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, 1(1–2):71–83. (Cited on pages 5, 9, and 16.)

Ferretti, T. R., Rohde, H., Kehler, A., and Crutchley, M. (2009). Verb aspect, event structure, and coreferential processing. *Journal of Memory and Language*, 61(2):191–205. (Cited on pages 76, 78, and 100.)

Frazier, L., Taft, L., Roeper, T., Clifton, Jr., C., and Ehrlich, K. (1984). Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition*, 12(5):421–430. (Cited on page 44.)

Friedrich, C. K., Alter, K., and Kotz, S. A. (2001). An electrophysiological response to different pitch contours in words. *NeuroReport*, 12(15):3189–3191. (Cited on pages 80 and 104.)

Garnham, A., Oakhill, J., and Cruttenden, H. (1992). The role of implicit causality and gender cue in the interpretation of pronouns. *Language and Cognitive Processes*, 7(3–4):231–255. (Cited on page 36.)

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes*, 23(3):265–304. (Cited on page 18.)

Gernsbacher, M. A., Hargreaves, D. J., and Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language*, 28(6):735–755. (Cited on page 20.)

Gioia, G. A., Isquith, P. K., Kenworthy, L., and Barton, R. M. (2002). Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology*, 8(2):121–137. (Cited on page 18.)

Grober, E. H., Beardsley, W., and Caramazza, A. (1978). Parallel function strategy in pronoun assignment. *Cognition*, 6(2):117–133. (Cited on page 44.)

Hagoort, P., Brown, C., and Groothusen, J. (1993). The syntactic positive shift (sps) as an erp measure of syntactic processing. *Language and Cognitive Processes*, 8(4):439–483. (Cited on page 78.)

Heim, S. and Alter, K. (2006). Prosodic pitch accents in language comprehension and production: ERP data and acoustic analyses. *Acta Neurobiologiae Experimentalis*, 66(1):55–68. (Cited on pages 80, 81, 91, and 104.)

Hendriks, P. and de Hoop, H. (2001). Optimality theoretic semantics. *Linguistics and Philosophy*, 24(1):1–32. (Cited on page 109.)

Hirschberg, J. and Ward, G. (1991). Accent and bound anaphora. *Cognitive Linguistics*, 2(2):101–121. (Cited on pages 41, 46, and 51.)

Hirst, W. and Brill, G. A. (1980). Contextual aspects of pronoun assignment. *Journal of Verbal Learning and Verbal Behavior*, 19(2):168–175. (Cited on page 23.)

Hoeks, J. C. J., Brouwer, H., Hendriks, P., and Stowe, L. A. (n.d.). Questions left unanswered. How the brain responds to missing information. Manuscript submitted to Brain and Language. (Cited on page 100.)

Hoeks, J. C. J. and Hendriks, P. (2011). Optimality theory and human sentence processing: towards a cross-modular analysis of coordination. *Journal of Cognitive Science*, 12(1):83–128. (Cited on pages 6, 37, and 109.)

Hoeks, J. C. J., Stowe, L. A., and Doedens, G. (2004). Seeing words in context: the interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, 19(1):59–73. (Cited on page 78.)

Huang, Y. (1991). A neo-Gricean pragmatic theory of anaphora. *Journal of Linguistics*, 27(02):301–335. (Cited on pages 4, 40, and 46.)

Huang, Y. (2000). Discourse anaphora: Four theoretical models. *Journal of Pragmatics*, 32(2):151–176. (Cited on pages 2, 40, 45, 46, 61, 70, and 71.)

Jakobson, R. (1960). *Style in Language*, chapter Closing Statement: Linguistics and Poetics, pages 350–377. MIT Press, Cambridge, MA. (Cited on page 1.)

Järvikivi, J., van Gompel, G., R. P., Hyöna, J., and Bertram, R. (2005). Ambiguous pronoun resolution contrasting the first-mention and subject-preference accounts. *Psychological Science*, 16(4):260–264. (Cited on pages 2 and 70.)

Kaan, E. and Swaab, T. Y. (2003). Repair, revision and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15(1):98–110. (Cited on page 78.)

Kaiser, E. (2011). Focusing on pronouns: consequences of subjecthood, pronominalisation, and contrastive focus. *Language and Cognitive Processes*, 26(10):1625–1666. (Cited on page 104.)

Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar*. CSLI publications Stanford, CA. (Cited on pages 2, 49, 61, and 70.)

Kehler, A., Kertz, L., Rohde, H., and Elman, J. (2008). Coherence and coreference revisited. *Journal of Semantics*, 25(1):1–44. (Cited on pages 4, 40, 44, and 45.)

Koornneef, A. W. and Sanders, T. J. (2012). Establishing coherence relations in discourse: the influence of implicit causality and connectives on pronoun resolution. *Language & Cognitive Processes*, ahead-of-print:1–38. (Cited on page 100.)

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, 1146(1):23–49. (Cited on pages 78, 79, and 80.)

Kuperberg, G. R., Paczynski, M., and Ditman, T. (2011). Establishing causal coherence across sentences: An ERP study. *Journal of Cognitive Neuroscience*, 23(5):1230–1246. (Cited on page 102.)

Kutas, M. and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the n400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62:621–647. (Cited on page 101.)

Kutas, M. and Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947):161–163. (Cited on page 77.)

Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press. (Cited on page 3.)

Lakoff, G. (1971). Presupposition and relative well-formedness. In et al., A. R., editor, *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*, pages 329–340. Cambridge University Press, Cambridge, United Kingdom. (Cited on page 41.)

Lambrecht, K. (1996). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*, volume 71. Cambridge University Press, Cambridge, United Kingdom. (Cited on pages 2, 3, and 75.)

Lamers, M. and de Hoop, H. (2005). *Constraint Solving and Language Processing*, chapter Animacy information in human sentence processing: An incremental optimization of interpretation approach, pages 158–171. Springer, Heidelberg, Germany. (Cited on page 109.)

Lau, E., Stroud, C., Plesch, S., and Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, 98(1):74–88. (Cited on pages 100 and 101.)

Levinson, S. C. (1991). Pragmatic reduction of the binding conditions revisited. *Journal of Linguistics*, 27(01):107–161. (Cited on page 40.)

Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. MIT press. (Cited on pages 45 and 71.)

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. The MIT Press, Cambridge, MA. (Cited on page 8.)

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, 9(2):157–201. (Cited on page 13.)

MacDonald, M. C., Just, M. A., and Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, 24(4):56–98. (Cited on page 13.)

MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4):676–703. (Cited on page 13.)

Magne, C., Astésano, C., Lacheret-Dujour, A., Morel, M., Alter, K., and Besson, M. (2005). On-line processing of "pop-out" words in spoken french dialogues. *Journal of Cognitive Neuroscience*, 17(5):740–756. (Cited on pages 91 and 104.)

McRae, K., Spivey-Knowlton, M. J., and Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38(3):283–312. (Cited on pages 5, 9, and 13.)

Morrett, L. M. and MacWhinney, B. (2013). Syntactic transfer in english-speaking spanish learners. *Bilingualism: Language and Cognition*, 16(1):132–151. (Cited on pages 2 and 112.)

Nieuwland, M. S., Petersson, K. M., and Van Berkum, J. J. A. (2007). On sense and reference: examining the functional neuroanatomy of referential processing. *Neuroimage*, 37(3):993–1004. (Cited on pages 15 and 110.)

Nieuwland, M. S. and Van Berkum, J. J. A. (2006). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain Research*, 1118(1):155–167. (Cited on pages 77, 100, and 101.)

Nieuwland, M. S. and Van Berkum, J. J. A. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and Language*, 106:119–131. (Cited on page 91.)

Osterhout, L. and Holcomb, P. J. (1995). Event-related potentials and language comprehension. In Rugg, M. D. and Coles, M. G. H., editors, *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*, pages 171–215. Oxford University Press, New York, NY. (Cited on page 78.)

Osterhout, L. and Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34(6):739–773. (Cited on pages 76, 77, 91, 99, 100, and 103.)

Pickering, M. J. and Traxler, M. J. (1998). Plausibility and recovery from garden paths: an eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24:940–961. (Cited on page 15.)

Pickering, M. J., Traxler, M. J., and Crocker, M. W. (2000). Ambiguity in sentence processing: evidence against frequency-based accounts. *Journal of Memory and Language*, 43(3):447–475. (Cited on page 15.)

Pierrehumbert, J. and Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. M. and Pollack, M., editors, *Intentions in Communication*, pages 271–311. MIT Press, Cambridge MA. (Cited on page 3.)

Poesio, M., Ponzetto, S., and Versley, Y. (2011). Computational models of anaphora resolution: A survey. *Linguistic Issues in Language Technology*. (Cited on page 40.)

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 118(10). (Cited on page 80.)

Prince, E. F. (1981). *Radical Pragmatics*, chapter Toward a taxonomy of given-new information, pages 223–255. Academic Press, New York, NY. (Cited on page 2.)

Roth, R. M., Isquith, P. K., and Gioia, G. A. (2005). *BRIEF-A: Behavior Rating Inventory of Executive Function–Adult Version: Professional manual.* Psychological Assessment Resources, Lutz, FL,. (Cited on page 19.)

Sanford, A. J. and Sturt, P. (2002). Depth of processing in language comprehension: not noticing the evidence. *TRENDS in Cognitive Sciences*, 6(9):382–386. (Cited on page 17.)

Schumacher, P. B. and Baumann, S. (2010). Pitch accent type affects the N400 during referential processing. *NeuroReport*, 21(9):618–622. (Cited on page 91.)

Smallwood, J. (2007). The lights are on but no one's home: meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin and Review*, 14(3):527–533. (Cited on page 18.)

Smallwood, J., McSpadden, M., and Schooler, J. W. (2008). When attention matters: the curious incident of the wandering mind. *Memory and Cognition*, 36(6):1144–1150. (Cited on page 18.)

Smallwood, J. and Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, 132(6):946–958. (Cited on page 18.)

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, 23(3):197–229. (Cited on pages 4, 40, 44, and 45.)

Spivey, M. (2008). *The Continuity of Mind*. Oxford University Press, Oxford, United Kingdom. (Cited on pages 6 and 37.)

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., and Sedivy, J. C. (2002). Eye movements and spoken language comprehension: effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, 45(4):447–481. (Cited on page 109.)

Steedman, M. (2007). Information-structural semantics for English intonation. In *Topic and Focus*, pages 245–264. Springer, Dordrecht, The Netherlands. (Cited on page 3.)

Streb, J., Rösler, F., and Hennighausen, E. (1999). Event-related responses to pronoun and proper name anaphors in parallel and nonparallel discourse structures. *Brain and Language*, 70(2):273–286. (Cited on pages 76 and 78.)

Swets, B., Desmet, T., Clifton, Jr., C., and Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory and Cognition*, 36(1):201–216. (Cited on pages 5, 12, 16, 17, and 107.)

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., and Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217):1632–1634. (Cited on page 38.)

Tavano, E. and Kaiser, E. (2008). Effects of stress and coherence on pronoun interpretation. In *Poster presented at the 21st annual CUNY Conference on Human Sentence Processing*, pages 13–15, University of North Carolina-Chapel Hill. (Cited on page 41.)

Taylor, R., Stowe, L. A., Redeker, G., and Hoeks, J. C. J. (2013). Comprehension of marked pronouns in Spanish and English: object anaphors cross-linguistically. *Quarterly Journal of Experimental Psychology*, (ahead-of-print):1–61. (Cited on pages 39, 74, 103, and 104.)

Traxler, M. J., Pickering, M. J., Clifton, Jr., C., and Van Gompel, R. (2000). *Evaluating theories of modification of complex noun phrases*, pages 149–174. Kluwer, Dordrecht, The Netherlands. (Cited on page 15.)

Trueswell, J. C. and Tanenhaus, M. K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In Clifton, Jr., C., Frazier, L., and Rayner, K., editors, *Perspectives in Sentence Processing*, pages 155–179. Lawrence Earlbaum Associates, Hillsdale, NJ. (Cited on page 13.)

Vallduví, E. and Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics*, 34(3):459–520. (Cited on page 43.)

Van Berkum, J. J. A., Brown, C. M., and Hagoort, P. (1999). Early referential context effects in sentence processing: evidence from event-related brain potentials. *Journal of Memory and Language*, 41(2):147–182. (Cited on page 78.)

Van Berkum, J. J. A., Brown, C. M., and Hagoort, P. (1999). When does gender constrain parsing? evidence from erps. *Journal of Psycholinguistic Research*, 28(5):555–566. (Cited on page 102.)

Van Berkum, J. J. A., Koornneef, A. W., Otten, M., and Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, 1146:158–171. (Cited on pages 8, 15, 76, 78, 100, 109, and 110.)

Van den Noort, M., Bosch, P., Haverkort, M., and Hugdahl, K. (2008). A standard computerized version of the reading span

test in different languages. *European Journal of Psychological Assessment*, 24(1):35–42. (Cited on page 18.)

Van Gompel, R. P. G., Pickering, M. J., Pearson, J., and Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, 52(2):284–307. (Cited on page 15.)

Van Gompel, R. P. G., Pickering, M. J., and Traxler, M. J. (2000). Unrestricted race: A new model of syntactic ambiguity resolution. In Kennedy, A., Radach, R., Heller, D., and Pynte, J., editors, *Reading as a Perceptual Process*, pages 621–648. Elsevier, Oxford. (Cited on pages 5, 9, and 15.)

Van Gompel, R. P. G., Pickering, M. J., and Traxler, M. J. (2001). Reanalysis in sentence processing: evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, 45:225–258. (Cited on pages 15 and 17.)

Van Rij, J., van Rijn, H., and Hendriks, P. (2013). How WM load influences linguistic processing in adults: a computational model of pronoun interpretation in discourse. *Topics in Cognitive Science*. (Cited on pages 18 and 108.)

Venditti, J. J., Stone, M., Nanda, P., and Tepper, P. (2001). Toward an account of accented pronoun interpretation in discourse context: Evidence from eye-tracking. *Rutgers Center for Cognitive Science: Technical Report RuCCSTR-68*. (Cited on pages 7, 41, and 44.)

Venditti, J. J., Stone, M., Nanda, P., and Tepper, P. (2002). Discourse constraints on the interpretation of nuclear-accented pronouns. In *Speech Prosody 2002, International Conference*. (Cited on page 104.)

Wang, L., Bastiaansen, M., Yang, Y., and Hagoort, P. (2011). The influence of information structure on the depth of semantic processing: How focus and pitch accent determine the size of the N400 effect. *Neuropsychologia*, 49(5):813–820. (Cited on pages 71, 75, 80, and 100.)

Ward, G., Birner, B., and Kaiser, E. (n.d.). The Oxford Handbook of Pragmatics. To be printed in 2014. (Cited on pages 2, 3, and 75.)

Waters, G. S. and Caplan, D. (1996). The capacity theory of sentence comprehension: critique of Just and Carpenter (1992). *Psychological Review*, 103(4):761–772. (Cited on page 18.)

Watson, D. G., Tanenhaus, M. K., and Gunlogson, C. A. (2008). Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science*, 32(7):1232–1244. (Cited on page 51.)

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

http://code.google.com/p/classicthesis/

Final Version as of June 7, 2013 (`classicthesis` version 4.1).

GRONINGEN DISSERTATIONS IN LINGUISTICS (GRODIL)

1. Henriëtte de Swart (1991). *Adverbs of Quantification: A Generalized Quantifier Approach.*
2. Eric Hoekstra (1991). *Licensing Conditions on Phrase Structure.*
3. Dicky Gilbers (1992). *Phonological Networks. A Theory of Segment Representation.*
4. Helen de Hoop (1992). *Case Configuration and Noun Phrase Interpretation.*
5. Gosse Bouma (1993). *Nonmonotonicity and Categorial Unification Grammar.*
6. Peter Blok (1993). *The Interpretation of Focus: an epistemic approach to pragmatics.*
7. Roelien Bastiaanse (1993). *Studies in Aphasia.*
8. Bert Bos (1993). *Rapid User Interface Development with the Script Language Gist.*
9. Wim Kosmeijer (1993). *Barriers and Licensing.*
10. Jan-Wouter Zwart (1993). *Dutch Syntax: A Minimalist Approach.*
11. Mark Kas (1993). *Essays on Boolean Functions and Negative Polarity.*
12. Ton van der Wouden (1994). *Negative Contexts.*
13. Joop Houtman (1994). *Coordination and Constituency: A Study in Categorial Grammar.*
14. Petra Hendriks (1995). *Comparatives and Categorial Grammar.*
15. Maarten de Wind (1995). *Inversion in French.*
16. Jelly Julia de Jong (1996). *The Case of Bound Pronouns in Peripheral Romance.*
17. Sjoukje van der Wal (1996). *Negative Polarity Items and Negation: Tandem Acquisition.*
18. Anastasia Giannakidou (1997). *The Landscape of Polarity Items.*
19. Karen Lattewitz (1997). *Adjacency in Dutch and German.*
20. Edith Kaan (1997). *Processing Subject-Object Ambiguities in Dutch.*
21. Henny Klein (1997). *Adverbs of Degree in Dutch.*
22. Leonie Bosveld-de Smet (1998). *On Mass and Plural Quantification: The Case of French 'des'/'du'-NPs.*
23. Rita Landeweerd (1998). *Discourse Semantics of Perspective and Temporal Structure.*

24. Mettina Veenstra (1998). *Formalizing the Minimalist Program.*
25. Roel Jonkers (1998). *Comprehension and Production of Verbs in Aphasic Speakers.*
26. Erik F. Tjong Kim Sang (1998). *Machine Learning of Phonotactics.*
27. Paulien Rijkhoek (1998). *On Degree Phrases and Result Clauses.*
28. Jan de Jong (1999). *Specific Language Impairment in Dutch: Inflectional Morphology and Argument Structure.*
29. Hae-Kyung Wee (1999). *Definite Focus.*
30. Eun-Hee Lee (2000). *Dynamic and Stative Information in Temporal Reasoning: Korean Tense and Aspect in Discourse.*
31. Ivilin Stoianov (2001). *Connectionist Lexical Processing.*
32. Klarien van der Linde (2001). *Sonority Substitutions.*
33. Monique Lamers (2001). *Sentence Processing: Using Syntactic, Semantic, and Thematic Information.*
34. Shalom Zuckerman (2001). *The Acquisition of "Optional" Movement.*
35. Rob Koeling (2001). *Dialogue-Based Disambiguation: Using Dialogue Status to Improve Speech Understanding.*
36. Esther Ruigendijk (2002). *Case Assignment in Agrammatism: a Cross-linguistic Study.*
37. Tony Mullen (2002). *An Investigation into Compositional Features and Feature Merging for Maximum Entropy-Based Parse Selection.*
38. Nanette Bienfait (2002). *Grammatica-onderwijs aan allochtone jongeren.*
39. Dirk-Bart den Ouden (2002). *Phonology in Aphasia: Syllables and Segments in Level-specific Deficits.*
40. Rienk Withaar (2002). *The Role of the Phonological Loop in Sentence Comprehension.*
41. Kim Sauter (2002). *Transfer and Access to Universal Grammar in Adult Second Language Acquisition.*
42. Laura Sabourin (2003). *Grammatical Gender and Second Language Processing: An ERP Study.*
43. Hein van Schie (2003). *Visual Semantics.*
44. Lilia Schürcks-Grozeva (2003). *Binding and Bulgarian.*
45. Stasinos Konstantopoulos (2003). *Using ILP to Learn Local Linguistic Structures.*
46. Wilbert Heeringa (2004). *Measuring Dialect Pronunciation Differences using Levenshtein Distance.*
47. Wouter Jansen (2004). *Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology Approach to English, Hungarian and Dutch.*
48. Judith Rispens (2004). *Syntactic and Phonological Processing in Developmental Dyslexia.*

49. Danielle Bougaïré (2004). *L'approche communicative des campagnes de sensibilisation en santé publique au Burkina Faso: les cas de la planification familiale, du sida et de l'excision.*
50. Tanja Gaustad (2004). *Linguistic Knowledge and Word Sense Disambiguation.*
51. Susanne Schoof (2004). *An HPSG Account of Nonfinite Verbal Complements in Latin.*
52. M. Begoña Villada Moirón (2005). *Data-driven identification of fixed expressions and their modifiability.*
53. Robbert Prins (2005). *Finite-State Pre-Processing for Natural Language Analysis.*
54. Leonoor van der Beek (2005). *Topics in Corpus-Based Dutch Syntax.*
55. Keiko Yoshioka (2005). *Linguistic and gestural introduction and tracking of referents in L1 and L2 discourse.*
56. Sible Andringa (2005). *Form-focused instruction and the development of second language proficiency.*
57. Joanneke Prenger (2005). *Taal telt! Een onderzoek naar de rol van taalvaardigheid en tekstbegrip in het realistisch wiskundeonderwijs.*
58. Neslihan Kansu-Yetkiner (2006). *Blood, Shame and Fear: Self-Presentation Strategies of Turkish Women's Talk about their Health and Sexuality.*
59. Mónika Z. Zempléni (2006). *Functional imaging of the hemispheric contribution to language processing.*
60. Maartje Schreuder (2006). *Prosodic Processes in Language and Music.*
61. Hidetoshi Shiraishi (2006). *Topics in Nivkh Phonology.*
62. Tamás Biró (2006). *Finding the Right Words: Implementing Optimality Theory with Simulated Annealing.*
63. Dieuwke de Goede (2006). *Verbs in Spoken Sentence Processing: Unraveling the Activation Pattern of the Matrix Verb.*
64. Eleonora Rossi (2007). *Clitic production in Italian agrammatism.*
65. Holger Hopp (2007). *Ultimate Attainment at the Interfaces in Second Language Acquisition: Grammar and Processing.*
66. Gerlof Bouma (2008). *Starting a Sentence in Dutch: A corpus study of subject- and object-fronting.*
67. Julia Klitsch (2008). *Open your eyes and listen carefully. Auditory and audiovisual speech perception and the McGurk effect in Dutch speakers with and without aphasia.*
68. Janneke ter Beek (2008). *Restructuring and Infinitival Complements in Dutch.*
69. Jori Mur (2008). *Off-line Answer Extraction for Question Answering.*
70. Lonneke van der Plas (2008). *Automatic Lexico-Semantic Acquisition for Question Answering.*

71. Arjen Versloot (2008). *Mechanisms of Language Change: Vowel reduction in 15th century West Frisian.*
72. Ismail Fahmi (2009). *Automatic term and Relation Extraction for Medical Question Answering System.*
73. Tuba Yarbay Duman (2009). *Turkish Agrammatic Aphasia: Word Order, Time Reference and Case.*
74. Maria Trofimova (2009). *Case Assignment by Prepositions in Russian Aphasia.*
75. Rasmus Steinkrauss (2009). *Frequency and Function in WH Question Acquisition. A Usage-Based Case Study of German L1 Acquisition.*
76. Marjolein Deunk (2009). *Discourse Practices in Preschool. Young Children's Participation in Everyday Classroom Activities.*
77. Sake Jager (2009). *Towards ICT-Integrated Language Learning: Developing an Implementation Framework in terms of Pedagogy, Technology and Environment.*
78. Francisco Dellatorre Borges (2010). *Parse Selection with Support Vector Machines.*
79. Geoffrey Andogah (2010). *Geographically Constrained Information Retrieval.*
80. Jacqueline van Kruiningen (2010). *Onderwijsontwerp als conversatie. Probleemoplossing in interprofessioneel overleg.*
81. Robert G. Shackleton (2010). *Quantitative Assessment of English-American Speech Relationships.*
82. Tim Van de Cruys (2010). *Mining for Meaning: The Extraction of Lexico-semantic Knowledge from Text.*
83. Therese Leinonen (2010). *An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects.*
84. Erik-Jan Smits (2010). *Acquiring Quantification. How Children Use Semantics and Pragmatics to Constrain Meaning.*
85. Tal Caspi (2010). *A Dynamic Perspective on Second Language Development.*
86. Teodora Mehotcheva (2010). *After the fiesta is over. Foreign language attrition of Spanish in Dutch and German Erasmus Students.*
87. Xiaoyan Xu (2010). *English language attrition and retention in Chinese and Dutch university students.*
88. Jelena Prokić (2010). *Families and Resemblances.*
89. Radek Šimík (2011). *Modal existential wh-constructions.*
90. Katrien Colman (2011). *Behavioral and neuroimaging studies on language processing in Dutch speakers with Parkinson's disease.*
91. Siti Mina Tamah (2011). *A Study on Student Interaction in the Implementation of the Jigsaw Technique in Language Teaching.*

92. Aletta Kwant (2011). *Geraakt door prentenboeken. Effecten van het gebruik van prentenboeken op de sociaal-emotionele ontwikkeling van kleuters.*
93. Marlies Kluck (2011). *Sentence amalgamation.*
94. Anja Schüppert (2011). *Origin of asymmetry: Mutual intelligibility of spoken Danish and Swedish.*
95. Peter Nabende (2011). *Applying Dynamic Bayesian Networks in Transliteration Detection and Generation.*
96. Barbara Plank (2011). *Domain Adaptation for Parsing.*
97. Çağrı Çöltekin (2011). *Catching Words in a Stream of Speech: Computational simulations of segmenting transcribed child-directed speech.*
98. Dörte Hessler (2011). *Audiovisual Processing in Aphasic and Non-Brain-Damaged Listeners: The Whole is More than the Sum of its Parts.*
99. Herman Heringa (2012). *Appositional constructions.*
100. Diana Dimitrova (2012). *Neural Correlates of Prosody and Information Structure.*
101. Harwintha Anjarningsih (2012). *Time Reference in Standard Indonesian Agrammatic Aphasia.*
102. Myrte Gosen (2012). *Tracing learning in interaction. An analysis of shared reading of picture books at kindergarten.*
103. Martijn Wieling (2012). *A Quantitative Approach to Social and Geographical Dialect Variation.*
104. Gisi Cannizzaro (2012). *Early word order and animacy.*
105. Kostadin Cholakov (2012). *Lexical Acquisition for Computational Grammars. A Unified Model.*
106. Karin Beijering (2012). *Expressions of epistemic modality in Mainland Scandinavian. A study into the lexicalization-grammaticalization-pragmaticalization interface.*
107. Veerle Baaijen (2012). *The development of understanding through writing.*
108. Jacolien van Rij (2012). *Pronoun processing: Computational, behavioral, and psychophysiological studies in children and adults.*
109. Ankelien Schippers (2012). *Variation and change in Germanic long-distance dependencies.*
110. Hanneke Loerts (2012). *Uncommon gender: Eyes and brains, native and second language learners, & grammatical gender.*
111. Marjoleine Sloos (2013). *Frequency and phonological grammar: An integrated approach. Evidence from German, Indonesian, and Japanese.*
112. Aysa Arylova (2013). *Possession in the Russian clause. Towards dynamicity in syntax.*
113. Daniël de Kok (2013). *Reversible Stochastic Attribute-Value Grammars.*

114. Gideon Kotzé (2013). *Complementary approaches to tree alignment: Combining statistical and rule-based methods.*
115. Fridah Katushemererwe (2013). *Computational Morphology and Bantu Language Learning: an Implementation for Runyakitara.*

Grodil
Secretary of the Department of General Linguistics
Postbus 716
9700 AS Groningen
The Netherlands